



An overview of current knowledge about the impacts of forest management certification

A proposed framework for its evaluation

Claudia Romero

Francis E. Putz

Manuel R. Guariguata

Erin O. Sills

Paolo O. Cerutti

Guillaume Lescuyer



An overview of current knowledge about the impacts of forest management certification

A proposed framework for its evaluation

Claudia Romero

Francis E. Putz

Manuel R. Guariguata

Erin O. Sills

Paolo O. Cerutti

Guillaume Lescuyer

Occasional Paper 91

© 2013 Center for International Forestry Research
All rights reserved

ISBN 978-602-1504-06-2

Romero, C., Putz, FE., Guariguata, MR., Sills, EO., Cerutti, PO. and Lescuyer, G. 2013. An overview of current knowledge about the impacts of forest management certification: A proposed framework for its evaluation. Occasional Paper 91. CIFOR, Bogor, Indonesia.

Photo by Michael Padmanaba/CIFOR

Certified timber in logs pond in PT. Sumalindo Lestari Jaya 2, West Kutai district, East Kalimantan, Indonesia. Timber certification is one mechanism for ensuring sustainable forest management.

CIFOR
Jl. CIFOR, Situ Gede
Bogor Barat 16115
Indonesia

T +62 (251) 8622-622
F +62 (251) 8622-100
E cifor@cgiar.org

cifor.org

Any views expressed in this book are those of the authors. They do not necessarily represent the views of CIFOR, the editors, the authors' institutions, the financial sponsors or the reviewers.

Table of contents

Abbreviations	v
Acknowledgements	vi
1 Introduction	1
Diversity of forest management certification schemes	3
Synergies with other policies related to forest governance	5
2 Evaluation in the context of forest certification	6
Background on the evaluation of conservation interventions	7
Previous assessments of forest certification impacts	10
3 Towards an improved framework for the evaluation of forest certification	12
A proposed roadmap	12
4 Knowledge needs for the design of an evaluation of the impacts of forest management certification	20
Knowledge about the forest sector	20
Knowledge about the temporal dynamics of certification	20
Knowledge about the drivers of self-selection	22
Knowledge about the political economy of the forestry sector	25
Knowledge about land use change	25
4 Concluding remarks	27
6 References	28

List of figures, tables and boxes

Figures

1	Extent of natural forest (in ha) certified by the Forest Stewardship Council (FSC) as of March 2013	4
2	Complementarity of the impact and process evaluations of certification that together facilitate the generation of new knowledge in a participatory manner (i.e., social learning) and frame adaptive management of forests	13
3	Example of a specific theory-of-change that accounts for reduction of work-related injuries in a certified FMU	17
4	The certification continuum	22
5	Contextual and other factors that influence FMUs' management decisions	23

Tables

1	National standards for forest management certification in tropical countries	3
2	Potential advantages and pitfalls of approaches for understanding the impacts of certification	8
3	Examples of expected outcomes of certification to be examined to determine the impacts of forest management certification	16
4	Risks and assumptions that underlie the transitions between outputs into short-, medium- and long-term outcomes from the forest certification intervention	18
5	Variables that affect forest management	21
6	Typology of attitudes regarding forest management	24

Boxes

1	The Forest Stewardship Council – FSC	2
2	FSC certification and evaluation approaches	15

Abbreviations

3ie	International Initiative for Impact Evaluation
ASI	Accreditation Services International
BMP	Best Management Practice
CAR	Corrective Action Request
CB	Certifying Body
DFID	Department for International Development
FLO	Fairtrade Labelling Organization
FSC	Forest Stewardship Council
FLEGT	Forest Law Enforcement, Governance and Trade
FMUs	Forest Management Units. Forests managed for timber purposes by private forest owners, communities, concessionaires, industrial groups and states legally recognized by corresponding authorities
FPIC	Free Prior and Informed Consent
HCVF	High Conservation Value Forest
IEG	Independent Evaluation Group – World Bank
IIED	International Institute for Environment and Development
ITTO	International Tropical Timber Organization
PEFC	Programme for the Endorsement of Certification
PES	Payments for Environmental Services
REDD+	Reducing Emissions from Deforestation and Forest Degradation, including conservation, sustainable management of forests and enhancement of forest carbon stocks
RIL	Reduced-Impact Logging
SCR	Steering Committee Report of the State-of-Knowledge Assessment of Standards and Certification
SFI	Sustainable Forestry Initiative
SLIMF	Small- and Low-Intensity Managed Forest
TFF	Tropical Forest Foundation
WWF	World Wildlife Fund

Acknowledgements

Thanks are due to David Raitzer for insightful observations and comments on a previous version. Also to Tuukka Castren and Peter Dewees (World Bank-Profor), Marion Karmann (Forest Stewardship Council) and Guntars Laguns (Accreditation Services International) for their assistance and/or various types of support during the preparation of this paper. Many of the ideas presented in this paper have also benefited with input from, and discussions with, Sylvie Gourlet-Fleury, Plinio Sist and Alain Karsenty (Centre de coopération internationale en recherche agronomique pour le développement – CIRAD). We also thank all the participants at a discussion workshop held in Montpellier, France, hosted by CIRAD during November 2011.

1 Introduction

Forest management certification is a market-based mechanism to promote sustainable use of forest resources. It recognizes responsible management through independently verified compliance with a set of underlying principles, criteria and indicators that delineate the ecological, social, economic and policy impacts resulting from forest management for specific objectives. As such, a credible label of certification makes the positive externalities of proper forest management visible to the public (Roberts 2012). The emergence of certification in the late 1980s was motivated by failures of other efforts to halt deforestation and improve forest management. While the launch of the ‘Forest Principles’ at the United Nations Rio Summit in 1992 recognized the urgency to manage forests to meet the needs of present and future generations, the global community could not come up with a legally binding agreement to halt forest loss, particularly in the tropics. Alarm about tropical forest destruction and concerns about the unintended consequences of boycotts of forest products inspired diverse stakeholders to collaborate on an initiative based on the concept of certification of forest management (Viana *et al.* 1996; Elliott 2000; Cashore *et al.* 2004; Auld *et al.* 2008; Cashore and Auld 2012; SCR 2012).

As is the case with many forest management and conservation interventions (e.g., payments for environmental services or establishment of protected areas), there is insufficient empirical evidence on the impacts of certification to generate lessons learned at the global scale. While several published reviews of forest management certification provide some guidance for future work, most were based on geographically limited case studies, indirect information or anecdotal observations and were not conducted by independent observers. Many forest stakeholders now agree on the need to critically assess **when, where, how, to what extent, why, at what cost to whom and for how long** certification

changed the ways forests are managed. In support of such assessments, this paper provides a critical description of known impacts of certification on the fates of natural forests in developing countries, along with a brief review of the literature on the evaluation of conservation interventions. Based on this background, the paper then proposes a roadmap towards the design of a formal evaluation of the empirical biophysical, social, economic and policy impacts of the Forest Stewardship Council (FSC) certification of timber extraction in natural tropical forests (Box 1). The paper also stresses that for a proper evaluation it is critical to understand the national and local contexts (social, political, biophysical, economic) that affect the implementation and ultimately permanence of certification’s impacts in a given forest. Much of the information needed to operationalize the roadmap presented here is considered essential to this end.

In this paper, ‘certification impacts’ refer to those changes in the forest itself and surrounding areas that are attributable to certification at several levels: neighboring local communities and workers; participating forest management units (FMUs), which are forests managed for timber production by private forest owners, concessionaires, industrial groups and states legally recognized by corresponding authorities; and, local and national governments and legal frameworks. Although assessing impacts along the forest product market chain (e.g., chain-of-custody auditing process) is important, the paper focuses on the forests and associated institutions and agents. It applies the definition of impacts of the Organization for Economic Co-operation and Development (OECD): ‘the positive and negative, primary and secondary long-term effects produced by a development intervention, directly or indirectly, intended or unintended’ (OECD 2002). Further, the use of the term ‘sustainability’ in this document refers to one of the goals of forest management certification; the more restricted

Box 1. The Forest Stewardship Council – FSC

The Forest Stewardship Council is an independent non-governmental organization that was created in 1993 in response to the need to curb forest loss and to incentivize sound forest management. A coalition among the WWF and several UK-based institutions followed-up on the failed dialogues at the Rio Summit in 1992 and helped develop what has become the most prominent certification scheme in tropical countries. The creation of the FSC involved a long deliberative process that involved representatives from environmental groups, social welfare groups and forest industries from the global North and South. Its structure and mechanisms facilitate participation of groups of stakeholders with often different interests and visions of what constitutes sustainable forest management and how that goal should be approached.

The FSC seeks to promote environmentally responsible, socially beneficial and economically viable management of the world's forests by establishing a labeling system associated with principles of ecological, social, economic and political sustainability. Stakeholders interested in forests and forestry participate in the FSC through institutional or individual voting memberships or as commentators on publically posted summaries of certification assessments. To level the playing field between the global North and South, representation in the social, environmental and business chambers is balanced.

FMUs interested in certification contact the FSC directly or through an FSC-accredited certifying body (CB), as illustrated in the table below. After the CB reviews the management plans and other documents prepared by the FMU, a pre-assessment visit by an interdisciplinary team from the CB will visit the FMU and interview other relevant stakeholders. On this basis, the CB then recommends which practices need to be improved and the corresponding timeframe before a full assessment is warranted (i.e., Corrective Action Requests: CARs). If the FMU makes the necessary changes and then contracts for, and passes, that audit, the CB recommends it for FSC certification. Unless valid objections to the CB's recommendation emerge during the public comment period, FSC certification is granted for five years pending annual surveillance by the CB. These annual visits from CB representatives may reveal further changes in management practices needed to comply with FSC Principles and Criteria. These CARs can be minor or major; failure to adequately address a major CAR during the specified period results in immediate certificate suspension.

CERTIFYING BODY	ACCREDITED SINCE
Bureau Veritas Certification	July 2005
Control Union Certifications	October 2005
GFA Consulting Group	June 2000
IMAFLOA (only Brazil)	
Institut für Marktökologie (IMO)	July 1998
QMI-SAI Global Assurance Services (QMI)	August 2008
Rainforest Alliance	July 1995
SCS Global Services	July 1995
SGS	July 1995
Soil Association Woodmark	July 1995

To ensure that CB activities do not threaten its credibility, the FSC set up in 2006 an independent body called Accreditation Services International (ASI). Annual inspections by ASI include witness audits in the field with the CB auditors, as well as office visits and document examination. If ASI audits detect serious problems, the CB can lose FSC accreditation.

term 'responsible' refers to what is actually certified. The question of sustainability of forest management has been the focus of governmental and inter-governmental processes that have put forth collections of criteria and indicators to promote or purport to guarantee sustainable forest management; these include the Montreal Process, a framework that defines indicators associated with sustainable forest management in non-tropical countries (Prabhu *et al.* 1998; Washburn and Block 2001; ITTO 2005; Pintér *et al.* 2012). The rationale behind certification standards certainly builds upon, and has learned from, these debates.

Diversity of forest management certification schemes

Numerous certification schemes operate at national levels in several countries. The two largest international certification bodies are the Programme for the Endorsement of Certification (PEFC) and the Forest Stewardship Council (FSC; Box 1). Created in 1993, the FSC operates through broadly discussed and agreed principles, criteria and indicators and serves as the oversight and certification-granting organization to which accredited certifying bodies (CBs) report on the implementation and compliance of its standards. FSC has worked with other forest management certification schemes such as Lembaga Ekolabel Indonesia (LEI), a non-profit multi-stakeholder organization that promotes just and sustainable forest management. The other large certification scheme – PEFC (formerly the Pan-European Forest Certification) works mostly in Europe and the USA and in a very different manner than the FSC. Most fundamentally, the PEFC has no principles and criteria of its own, but instead recognizes the activities of autonomous certification schemes. In tropical countries, these schemes include the Brazilian Forest Certification Program (CERFLOR); the Malaysian Timber Certification Council (MTCC); and the Pan-African Forest Certification (PAFC) in Gabon. In other words, PEFC provides another layer of assurance that the standards of approved certification bodies are met (SCR 2012).

To render FSC principles and criteria more operational and locally appropriate, most countries and regions have, or are developing, detailed sets of standards, indicators and verifiers against which representatives of certifying bodies audit (Table 1). To be eligible for FSC approval, these

national and regional standards must be developed in a participatory manner according to specified procedures. In the absence of detailed standards, CBs can propose for approval by the FSC their own sets of indicators based on FSC Principles and Criteria (FSC 2013) and, on some occasions, develop their own generic regional standards (e.g., Rainforest Alliance for Indonesia). The FSC is the largest certification scheme in the tropics (Figure 1). While this paper focuses on the evaluation of FSC certification, its approach could also apply to PEFC.

Numerous studies have compared forest certification schemes, particularly on the basis of the stringency of their standards, modes of operation and constituencies (Vogt *et al.* 2000; Cashore 2002; Holvoet and Muys 2003; Oliver 2004; Fischer *et al.* 2005; WWF/World Bank Global Forest Alliance 2006; Auld *et al.* 2008; McDermott *et al.* 2008; Tikina and Innes 2008; Overdevest 2010; Clark and Kozar 2011; Johansson and Gun 2011). Another study compared the standards of the FSC and the North American-based Sustainable Forestry Initiative (SFI) with the Montreal Process. Like certification, the Montreal Process aims to promote sustainable forest management, but it focuses at the national level and takes a descriptive approach. In contrast, FSC's forest management certification operates at the level of discrete or tightly linked FMUs in a prescriptive manner, with outcomes required (Washburn and Block 2001).

Table 1. National standards for forest management certification in tropical countries

COUNTRY	DATE
Bolivia	May 1998
Brazil *	July 2001
Cameroon	April 2012
Central African Republic ^a	April 2012
Colombia	February 2003
Democratic Republic of Congo ^a	April 2012
Ghana	July 2012
Mexico	December 2010
Peru	July 2001
Republic of Congo ^a	April 2012
Republic of Gabon ^{*a}	April 2012

a Congo Basin Regional Standard

Note: Some of these countries have also developed national certification schemes for forest management (indicated with an asterisk).

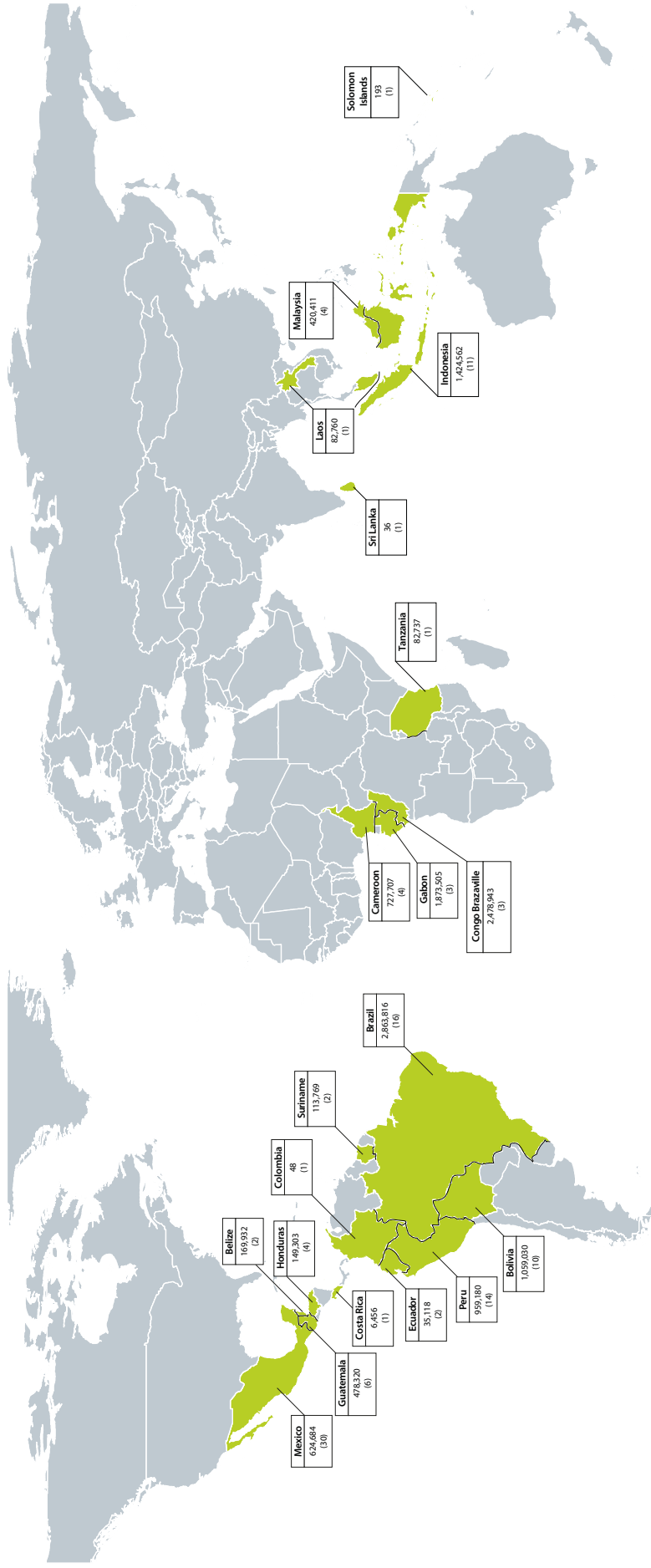


Figure 1. Extent of natural forest (in ha) certified by the Forest Stewardship Council (FSC) as of March 2013

Note: Number of certificates per country noted parenthetically.

Source: Forest Stewardship Council.

The proliferation of certification schemes is understandable given the broad scope of potential applications (national, regional or international), the diverse cultures of forest industries and their associated governmental agencies, and other biophysical and institutional differences in the places where these schemes have emerged. These certification schemes vary fundamentally in the involvement of stakeholders in defining their standards and in the degree to which their requirements exceed those set by national laws. They also vary in the breadth of their requirements. Overall, most of the comparative studies mentioned above concluded that the FSC is the strictest and most complete because it explicitly addresses sustainability related to:

- political issues (e.g., respect of applicable laws, FSC Principle 1);
- environmental issues (e.g., environmental impact, maintenance of high conservation value forests, FSC Principles 6 and 9);
- social issues (e.g., workers' rights and employment conditions; tenure, use rights and responsibilities; indigenous peoples' rights; community relations and benefits from the forests, FSC Principles 2, 3, 4 and 5); and
- economic issues (e.g., benefits from the forest, management plan, monitoring and assessment, FSC Principles 5, 7 and 8).

In addition, as Clark and Kozar (2011) and others point out, most of the other certification schemes focus only on parts of the production chain (e.g., legal, biophysical or social issues). It is also worth mentioning that proliferation of schemes and claims of sustainability of forest management remain a concern insofar as they can confuse consumers and jeopardize the credibility of certification (Putz 2004; Fischer *et al.* 2005; Eden 2009; Ruben and Zuñiga 2010; Treves and Jones 2010; Roberts 2012).

Other comparisons of certification schemes have recently explored governance issues (Marx and Cuypers 2011). One study in particular (McDermott *in press*) contrasted the FSC, PEFC, Fairtrade Labelling Organization (FLO) and the Climate, Community and Biodiversity Alliance. The author proposes a framework with three components associated with how decisions are made (e.g., who decides on what the standards should consider), how context determines participation in decision-making and how costs and benefits are distributed in order to analyse equity assurances in the contents of their standards.

Since 'equity' is a contested concept, the author explicitly discusses the goals and outcomes of equity as it pertains to these schemes. One finding is that stakeholders' level of trust in certifiers affects the certification process, and as a consequence, equity outcomes. For the FSC, the author reports that despite efforts to secure stakeholder participation in discussions of standards (e.g., empowering through engagement at several scales), equitable distribution of costs and benefits of certification were not addressed (i.e., stakeholder engagement does not guarantee equitable cost and benefits distribution).

Synergies with other policies related to forest governance

Forest certification does not act in a vacuum. It is implemented in particular social, institutional and political contexts that further influence decisions regarding forest use. One example is decentralization of control over resource management (e.g., Pacheco 2004; Agrawal 2007; Agrawal *et al.* 2008; Bowler *et al.* 2010; Brooks *et al.* 2012). Insights are also needed about how the impacts of certification compare with other interventions with shared goals including, most prominently, governmental regulations (Lee and Norris 2012). In addition, private forest companies that seek certification operate in countries that may be affected by various recent national and international efforts to assure the legality of forest products. These efforts could include, most prominently, the Forest Law Enforcement, Governance and Trade (FLEGT) action plan by the European Union and its Voluntary Partnership Agreements (www.eufl.rgt.int), as well as the amended Lacey Act (2008) in the USA (www.forestlegality.org). Although both of these initiatives focus on the legality of timber trade, they ultimately seek to foster better governance along the entire production chain – from harvesting to consumption. Synergies between efforts to assure legality and forest certification schemes may need to be explored to avoid redundancies (Carlsen *et al.* 2012; Tind 2012) and to otherwise increase their mutual effectiveness and efficiency (Vogel 2008). Furthermore, linkages between forest management certification and forest conservation interventions based on payments for environmental services (e.g., water, biodiversity and/or carbon) could enhance the effectiveness of certification and address some critical market failures in the timber trade (Hyde 2012).

2 Evaluation in the context of forest certification

As stated above, the need for a critical evaluation of the empirical impacts of forest certification has gone unfulfilled to date. Such impacts include: changes in the forest itself and surrounding areas; at the level of neighboring local communities and workers; for participating FMUs; and, to local and influenced institutions attributable to FSC certification. Among the many possible reasons for the lack of a critical evaluation are: the assumption that certification is inherently environmentally, economically, politically and socially beneficial; the cost of such a study; the fact that in some regions (e.g., the Congo Basin) certification only recently became important to forest management decisions; and, the inherent methodological and logistical challenges in evaluating the potential direct and indirect impacts of such a complex intervention in a wide variety of forests under a diverse set of ecological, socio-economic and political conditions. Additional constraints derive from the long-term scope of most conservation interventions and their often vague objectives (e.g., biodiversity conservation, maintenance of ecosystem integrity, sustainability or social well-being). This vagueness makes it difficult to identify specific elements to assess. Finally, in addition to lack of funding and incentives for evaluation is the lack of personnel trained in the rapidly evolving field of environmental impact evaluation (Ferraro 2009; Mickwitz and Birnbaum 2009). Lack of proper evaluation of the impacts of forest certification heightens the risk of both poor downward and upward accountability, including the process of certification itself (Rogers 2012).

It is also important to determine the extent to which certification delivers on its promises of maintaining forest values in order to inform its supporters if they are 'getting their money's worth' from the intervention. That is, understanding the factors and conditions that lead to impacts can help stakeholders compare the outcomes of certification to those returns from possible conservation investments (e.g., supporting green

markets as opposed to actively contributing to protected area management or other conservation alternatives). Further along the certification pathway, the tangible, less tangible and dynamic direct and indirect costs and benefits to FMUs of getting and remaining certified must be understood. In other words, the impacts of certification on the managed forests themselves and in neighboring areas need to be formally evaluated.

As an illustration of the consequences of the current lack of a proper evaluation of certification, imagine that certification of a particular FMU was associated with achievement of the biophysical objective of protecting riparian buffer zones. Unfortunately, buffer zone protection resulted in a conflict with a local community whose traditional uses of natural resources in these same areas were precluded. Awareness of such tradeoffs and other sorts of contested situations is clearly a first step towards their solution. It sets in motion deliberative processes and agreements involving all relevant parties, enforcement mechanisms, sanctions for lack of compliance and monitoring and verification activities performed in ways conducive to joint decision-making, cost minimization and benefit-sharing. Applying an evaluation lens at the onset of designing an intervention might help prevent such conflicts and, further down the road, facilitate identification of unexpected outcomes and suggest ways to deal with their occurrence.

In this paper we address two main goals of implementing a formal evaluation of forest certification: (i) to assess the extent to which certification is the direct and indirect driver of observed changes in the outcomes of forest management (i.e., the likely or achieved short- and medium-term effects of an intervention; OECD 2002); and (ii) to determine how other interventions and processes contribute to particular outcomes related to certification. As a whole, these hoped-for outcomes include maintenance or enhancement of forest values (e.g., biodiversity,

ecosystem service provision); social welfare of forest owners, workers and local people (e.g., health and education, access to credit, increased assets); the financial and legal status of certified FMUs; and changes in policy frameworks. To distinguish the impacts of certification from those related to other contextual factors and interventions [that is, to distinguish the impacts given in (i) and (ii) earlier in this paragraph], the causal pathways that led to any observed changes need to be understood. In short, the key factors that affect decisions driving the certification-related changes need to be identified (SCR 2012).

To identify the causal pathways and to understand how they operate requires a formal theory-of-change of the certification intervention (a theory-of-change is a model that describes how the process of change in itself occurs). For example, management decisions at the FMU level are made based on opportunities related to silvicultural knowledge, traditions, technical and financial capacities, market information, timber yields, inputs from social actors and governmental policies and regulations, among other factors. At the same time, decisions are constrained, or at least modulated by a range of factors including lack of technical capacity, as well as policy, institutional and market failures. The interplay of enabling and constraining factors clearly influences management choices and the resulting outcomes in the forest and further afield. That said, evaluations should be grounded in active participation by all relevant stakeholders to give jointly created knowledge credibility (i.e., true and technically adequate for handling evidence), salience (i.e., relevance and value to decision-makers and other evaluation users) and legitimacy (i.e., fairness of knowledge gathering, unbiased and respectful; Mollinga 2010; Rowe 2012). It should also be based on the shared belief that an entity's actions are desirable, proper and appropriate, based on institutionalized norms and practices that confer authority (Bernstein 2004). This emerging knowledge evolves as participants in the evaluation learn and as contexts change. Thus, the tools and processes of evaluation need to accommodate the new knowledge generated by participants and reflect new benchmarks in the collection of evidence of impacts (Mickwitz and Birnbaum 2009; Rogers 2009; van Stolk *et al.* 2011). Moreover, insufficient participation of appropriate stakeholders in the evaluation process hinders social learning and it may lead to equivocal

allocation of costs and benefits. In the absence of a mechanism for learning from past mistakes, other forest conservation interventions and the process of certification itself may suffer. In the particular case of certification, lack of a systematic assessment deprives stakeholders – from local social actors and resource managers to policy-makers and donors – of information to assess the tradeoffs among the different domains of forest management sustainability. Likewise, it denies them the knowledge needed to make evidence-based – and thus more informed – decisions.

To retain its utility, an evaluation of certification should also capture any changes that affect its impacts over time. Such changes could include reduced added-value of the certification intervention due to improved skills of both forest managers and administrators (e.g., harvesting costs diminished for the FMU due to streamlined forest operations and increased efficiency during timber harvesting). Furthermore, impacts that come from contextual changes beyond the certification intervention itself need to be addressed (e.g., low current additionality: early profits from certification might diminish over time with improved legal frameworks and enforcement, which would reduce the need for certification to maintain forest values).

Background on the evaluation of conservation interventions

Programme evaluation is a well-developed field, particularly in regards to assessing the impacts of public policies (e.g., welfare benefits such as improved health and education) and social and development interventions (e.g., conditional cash transfers). In contrast, evaluation of conservation interventions, including certification, remains a contested field of enquiry (Cook *et al.* 2010) and lags behind despite a recent flurry of publications (see later in this section). Evaluations seek to understand systems and processes through the generation of knowledge situated between research, policy and practice. As already mentioned, a main question is how an intervention changes specific variables (*attribution*) that produce particular outcomes. It is also important to discern how the ecological, social, economic and policy contexts in which an intervention is embedded influence its outcomes (*contribution*). That is, direct linear responses to certification and other complex interventions

are unlikely due to the multitude of natural and human-induced processes acting at different scales in both time and space (Levin *et al.* 2013). In these complex social–ecological systems, the multiple causal mechanisms that operate simultaneously are context dependent and prone to unpredictable feedback loops that give rise to emergent outcomes (i.e., recursive causality; Rogers 2008).

The indirect positive or negative effects of conservation interventions are particularly hard to capture. As one spillover effect, for example, an FMU might employ good management practices because it hired a worker formerly trained in a certified FMU rather than as a direct response to certification requirements. Nevertheless, these spillover effects need to be understood to assess the overall efficiency and effectiveness of conservation interventions such as certification (Nussbaum and Simula 2004). In the scenario above, evaluation of certification would require determining if certified FMUs are the only ones doing proper

management because they have hired all the high-quality available contractors (i.e., leakage). Likewise, evaluation processes should help establish if governments have focused on enforcement of uncertified FMUs because they believe that the tracking of compliance with legal requirements already takes place in certified units.

Several approaches have been used to evaluate conservation interventions. They include robust experimental methodologies through randomized control trials; more qualitative and less informative analyses on the causal effect of the intervention in the outcomes; and more qualitative non-experimental methods (Bamberger and White 2007; GAO 2009; Table 2). These approaches also include the now popular systematic reviews, for which studies with robust results are still much needed (Pullin *et al.* 2009; Bowler *et al.* 2010).

Positive self-selection

As with any voluntary scheme, participants self-select into certification. The resulting ‘positive

Table 2. Potential advantages and pitfalls of approaches for understanding the impacts of certification

METHOD	DESCRIPTION	LIMITATIONS
Experimental	Randomly selected FMUs were allocated to the forest certification intervention.	Selection bias is likely because certification is voluntary. A comparison based on the experimental approach is not feasible.
Quasi-experimental	Because the certification treatment was not randomly allocated, a comparison group of uncertified FMUs needs to be constructed (counterfactual). The treatment and control groups should only differ in their certification status.	Comparison group construction is data intensive and technically difficult. Approaches include matching techniques (e.g., groups of certified and non-certified FMUs matched by factors that influence certification outcomes) and instrumental variables (e.g., correlated and easier-to-assess variables are used to infer impacts), among others.
Before–after	Baseline data on key outcomes related to the certification intervention are measured and compared with data corresponding to the post-certification condition.	Data are often not available for all the variables before certification was granted for both treatment (i.e., certified) and control groups.
Systematic review	Intensive analyses of certified FMUs, drawing on the history of the FMU and how the particular nature of the mechanisms and contextual factors produced change.	Time-consuming and knowledge-demanding method: requires robust results of properly designed studies and thus fails to determine the integrated impacts of forest management certification unless available literature exists.
Expert judgment	Assess the impacts of certification through compilation and synthesis of statements of people with profound knowledge of certification and the contexts in which forest management occurs.	Because forest management certification is complex, this approach can be informative but may fail to capture the integrated effect of certification-driven changes and interactions with contextual factors.

Source: adapted from Romero and Castrén (2013)

Note: Table is based on the framework proposed by GAO (2009).

selection bias' can obscure insights that could be derived from random allocation of FMUs to the treatment (certified) and control (uncertified) groups, as in randomized experiments. Selection bias renders it more difficult to separate the effects of the certification intervention from the direct and indirect influences of other contextual factors; it also makes it more challenging to identify the extent to which the intervention caused any observed impacts (i.e., internal validity; Chen *et al.* 2011a). However, few studies of the impacts of conservation interventions adequately addressed selection bias. This oversight often resulted in exaggeration of the direct positive impacts of the interventions (Ferraro and Pattanayak 2006; Pattanayak *et al.*, 2010). These deficiencies have recently started to be addressed; several practical examples of ways to avoid them are now available for the conservation arena (Andam *et al.* 2008; Sims 2010; Ferraro *et al.* 2011; Nelson and Chomitz 2011; Alix-Garcia *et al.* 2012; Arriagada *et al.* 2012; Laufer *et al.* 2013).

A mixed bag of methodological approaches

Because the certification and other conservation interventions are typically not allocated at random, quasi-experimental approaches are often needed to ensure external validity (i.e., how the outcomes from the intervention could be generalized to fit other contexts; Chen *et al.* 2011a). Therefore, to gauge the intervention's impact, a counterfactual (i.e., the condition of those affected if the intervention had not occurred) needs to be constructed or selected (Greenstone and Gayer 2007; Ferraro 2009; Jagger *et al.* 2010). Because natural counterfactuals are rare (Rowe 2012), a comparison group usually needs to be constructed on the basis of variables expected to affect the outcomes of the intervention. A diversity of statistical techniques (e.g., instrumental variables – IV; regression discontinuity designs – RDD, matching) are used to analyse these situations. Partly because IVs are often difficult to identify and justify, matching is often used to construct a comparison group with similar distributions of variables (both observable and unobservable) that might affect both participation and outcomes (Rosenbaum and Rubin 1983). More recently, other approaches have been used where self-selection poses methodological problems (e.g., encouragement designs: Bradlow 1998; West *et al.* 2008; endogenous switching regression: Kleeman and Abdulai 2012).

In the case of certification, many variables that affect the maintenance of forest values are likely to also influence the probability of participation in certification. These variables could include FMU area, market orientation and institutional arrangements such as community based, private or public (for more examples see Table 2). Exact matching entails tradeoffs between the stringency of criteria used to select counterfactuals (e.g., the number of matching variables used) and the number of possible replicates. Quantitative studies on the impacts of conservation interventions should be complemented by qualitative insights gained through both a variety of techniques and available secondary information to triangulate evidence (i.e., addressing the same research question from different perspectives; Ton 2012). Also, governmental and other data are needed to establish the external validity of the certification intervention model.

Among the qualitative approaches to evaluation are exploratory and participatory methods. Exploratory methods (e.g., general elimination methodologies, process tracing and contribution analysis) can help explain what occurred and how and thereby help infer causality (Bamberger *et al.* 2009; Collier 2011; White and Phillips 2012). Participatory approaches rely on stakeholder perceptions to examine behavioral change, establish attribution and suggest how interventions can be improved (e.g., most significant change, the success case method, outcome mapping, stakeholder and multi-criteria analyses, participatory impact assessment and participatory social mapping; Rogers 2008; Chambers 2009; Rogers 2009; Schreckenberg *et al.* 2010; White and Phillips 2012). Qualitative approaches should still be based on a comparative framework such as before-and-after comparisons of both treated and control groups or interrupted time series (GAO 2009).

In summary, evaluating the impacts of forest certification requires a range of quantitative methods to reveal connections along the causal change model. These quantitative methods need to be complemented with qualitative approaches to detect the indirect effects of the intervention (i.e., a mixed-methods approach; Garbarino and Holland 2009; White 2009; Rugh *et al.* 2010; Bamberger 2012; Stern *et al.* 2012; Ton 2012).

Previous assessments of forest certification impacts

Forest certification has received considerable attention from a variety of groups concerned about the fates of the world's remaining forests. With few exceptions, previous attempts to assess the impacts of certification have focused on examining secondary information and stakeholder perceptions. Although none of the approaches reviewed below resemble what could be considered a well-designed evaluation, they nevertheless provide useful information and insights.

Interviews about the impacts of forest management certification

An indirect way to assess the impacts of certification is by soliciting impressions from representatives of FMUs, timber industries, local communities, government officials, environmental groups and buyers. Global and regional studies of this sort all reported that the compilers solicited impressions of both supporters and critics of forest certification but nevertheless reported mostly positive impressions of impacts (Frost *et al.* 2003; Hartsfield and Ostermeier 2003; Humphries and Kainer 2006; Ebeling and Yasué 2009; Sheil *et al.* 2010; Zagt *et al.* 2010). These studies all conclude that certification has done more to improve tropical forestry than any other global initiative (e.g., the Tropical Forestry Action Plan, ITTO Objective 2000). They also agree on the need for empirical assessments of certification impacts.

Literature reviews, systematic reviews and methodological constraints

Literature reviews on the impacts of forest certification have mostly focused on particular issues (e.g., biodiversity, wildlife, local communities) and regions. Typically, the reviews assessed one or a few specific aspects of management such as the protection of riparian buffer zones, seed-tree retention and deforestation. Among the diversity of approaches employed to determine changes in forest management due to certification, some studies employed before-and-after certification comparisons, while others compared certified and non-certified FMUs (e.g., Schulte-Herbruggen and Davies 2006; Karmann and Smith 2009). Other studies examined the apparent impacts of certification on different aspects of management, including community forestry (Molnar 2004; Corso *et al.* 2008), community enterprises (Butterfield *et al.* 2005), biodiversity (Walrecht *et al.* 2012) and changes in

forest cover as a function of the relative importance of the forest sector (e.g., high level of timber production and proportion of GDP; Damette and Delacote 2011).

To assess the current status of evaluation of certification, Blackman and Rivera (2010) analysed 134 documents on certification of timber, fish, bananas, coffee and general agricultural practices. Of these studies, only 14 employed designs that appropriately considered the confounding effects of selection bias. Of these 14, the sole study on forest management compared a certified and a nearby uncertified community forest operation in Brazil and reported that certification had small positive environmental and socio-economic impacts. Selection bias in this study was at least partially avoided because the communities had similar land tenure arrangements and both ran their own operations for the primary purpose of timber production (Barbosa de Lima *et al.* 2008). In a recent update of their 2010 study, Blackman and Rivera (2011) ranked 46 peer-reviewed studies on the impacts of certification on the basis of the robustness of their research designs. Although none of these studies was on timber certification, the results are nevertheless of interest. In 7 of the 11 best-designed studies, certification was not associated with increased environmental, social or economic benefits for the producers. The exceptions were for bananas, for which productivity increases led to on-farm investments; for coffee, for which the social and economic benefits of certification were due to particular conditions and did not translate into gains in education or health; and for tourism, for which certified operations received a price premium. Unfortunately, due to insufficient data on the individual practices that led to certification, formal meta-analyses were not feasible. It must be noted that although meta-analyses can be useful, they do not replace well-designed evaluations that can result in causal inferences about the contributions of the intervention to the maintenance of forest values and identify unintended and indirect impacts.

A more nuanced analytical approach to evaluating the impacts of certification that was recently tested is based on studies on compliance with best-management practices for forestry (BMPs; Newsom *et al.* 2012). The authors reviewed what they judged to be properly designed and well-executed studies on the impacts of set-

asides, including riparian buffer zones, on species diversity, population viability and the quality of water, air and soil. The authors reasoned that in the absence of properly designed evaluations of certification, adherence to BMPs that are also required for certification could provide insights into certification impacts. They contended that disaggregation of the activities required to achieve a desired outcome can aid in the assessment of certification impacts. For example, the observation that establishment of riparian buffer zones benefits sensitive taxa will inform hypotheses that address the underlying causal mechanisms. Disaggregation, also recommended by researchers for other evaluation challenges (Bamberger *et al.* 2009; Rogers 2009; Jagger *et al.* 2010), provides a basis for the evaluation of the empirical impacts of certification (Crosse *et al.* 2012). Unfortunately, disaggregation can also lead to disregard of the integrated and interacting range of impacts of the certification intervention *per se*. As such, this approach will miss the indirect effects and unintended impacts of certification because it is not based on a detailed theory-of-change.

Recently, the Steering Committee Report of the State-of-Knowledge Assessment of Standards and Certification thoroughly assessed the impacts of standards and certification for achievement of sustainability goals for agriculture, fisheries, forestry and aquaculture (SCR 2012). This effort was based on literature reviews, meetings with business leaders, interviews with key stakeholders and analyses of case studies. Through this lengthy and reflective process, the authors analyzed the contexts, actors, impacts, pathways and trends that will likely affect standards and certification systems in coming years. The resulting document is too rich to summarize in detail here, but it provides evidence of improvements in environmental, social and economic practices associated with, but not clearly attributable to, certification. In addition, the assessment reveals some unintended negative effects on a case-by-case basis. At larger scales in both space and time, the assessment concludes that it is hard to infer impacts of certification. Furthermore, the variety of methods and methodological limitations preclude identification of the causal effects of the certification intervention. The report goes so far as to suggest that the indirect impacts of certification might be of more consequence than direct ones. And its main conclusion is that the incentives available from certification will not be sufficient for FMUs

that employ far-from-acceptable management practices. The report also argues for the need to develop a robust evaluation of empirical impacts (i.e., from management practice to outcome) to properly assess the suitability of this private, voluntary, market-oriented policy to promote sustainability.

Review of corrective action requests (CARs)

Several assessments of forest certification are based on analyses of changes over time in corrective action requests (CARs) issued by FMU auditors from CBs. The FMUs are typically given deadlines to address each CAR; in the case of major CARs, failure to correct the identified problem within the specified time period results in refusal to grant the certificate, or else its immediate suspension. Reviewing the assignment and satisfaction of CARs provides insights into the nature of problematic management issues and their evolution, as well as indirect evidence of improved practices (Rametsteiner and Simula 2003; Newsom and Hewitt 2005; Newsom *et al.* 2006; Peña-Claros *et al.* 2009).

To date, assessments of certification impacts based on CARs have all reported a decline over time in the number of issues in need of correction during the auditing process; the authors of these assessments interpreted this decline as indirect evidence for positive impacts of certification. As the authors themselves made clear, such results need to be interpreted with caution due to the many limitations of using CARs as surrogates for more direct measures. One of these limitations is that FMUs may learn how to ‘manage for the audit’ so as to avoid sanctions but without actually correcting the problem. After all, auditors can spend little time in the field assessing management practices. Also, given that the business of forest certification depends on FMUs being certified, auditors could suffer from a perverse incentive to overlook problems with management practices (i.e., potential conflicts of interest; McDermott 2012). A more insidious problem could emerge if, to avoid certificate suspension, slightly different CARs were assigned to the same inherent problem during successive audits. It should also be noted that most of the changes in management practices indicated by tracking the evolution of CARs are based on process and not performance. This means the issued CARs are indirect, but of key importance, to the maintenance of forest values (Nussbaum and Simula 2004; Bartley 2007; Peña-Claros *et al.* 2009; van Kuijk *et al.* 2009).

3 Towards an improved framework for the evaluation of forest certification

As mentioned in the previous sections, evaluations of the impacts of certification need to consider differences among FMUs in biophysical, socio-economic and policy characteristics that affect how they are, and should be, managed. For example, forests vary in stocking of commercial species, terrain, accessibility, seasonality and underlying natural dynamics. Social aspects vary with the characteristics of communities living within the FMU's area of influence and their relationships with it (e.g., employment possibilities and freedom of access), including the understanding of community residents about the FMU and its operations. Governance aspects that influence forest management include tenure types and rules governing resource access, policies and regulations that define allowable harvests, required management procedures and fees (e.g., taxes and royalties), participation of local stakeholders and the extent of regulation enforcement (Coleman and Steed 2009; Burgess *et al.* 2012). Finally, characteristics of the economic sphere of influence on certification include FMU type (e.g., public, communal or private), firm size, markets accessed (e.g., international vs. local), type of harvesting arrangement (e.g., subcontracted or not), technology employed and features of the commercialization process (i.e., market chains and the extent of vertical integration; Amacher *et al.* 2009; Assunção *et al.* 2012). An effective evaluation strategy would have to consider this complexity and address associated methodological challenges. The proposed approach presented below is grounded on the intellectual foundations of evaluators experienced with complex interventions (e.g., P. Rogers, M. Bamberger, H. White, and M.Q. Patton); institutions that support debates on these issues (e.g., Independent Evaluation Group (IEG), Department for International Development (DFD), International Initiative for Impact Evaluation (3ie), Campbell Collaboration; Hivos E-Dialogues, betterevaluation.org); and several researchers (e.g., Jagger *et al.* 2010; Guijt *et al.* 2011; Wigboldus and Brouwers 2011; also see

references in Section 2). It also sits firmly on the foundation for adaptive management of natural resources formulated by C. S. Holling and his colleagues (Holling 1978), which is fundamental for the sustainable use of natural resources through continued learning and experimentation.

A proposed roadmap

Proposed key activities towards a well-informed evaluation of the empirical impacts of forest certification are outlined here. The first four activities are discussed in more detail below based on the review sections above. Ideas relevant to deliberative processes used to design the evaluation are also assembled. Activities 5 and beyond correspond to implementation of impact and process evaluations and then progress on to the activities that commence once the evaluation is concluded. Given current knowledge about the impacts of certification, it is not possible to discuss these activities in detail. Instead, in Section 4 key information is provided for the design phase that precedes the field-based evaluations. The steps are as follows:

1. **Clarify the values** that underpin the evaluation – what are the desired and undesired processes, impacts and distributions of costs and benefits for different types of stakeholders?
2. **Define the scope** or boundaries of the evaluation using a systems approach (Fujita 2010).
3. **Design parallel processes of:**
 - a. theory-based impact evaluation of the intervention (henceforth 'impact evaluation').
 - b. evaluation of the implementation process of the intervention (henceforth 'process evaluation').
4. **Identify initial questions to be addressed by the evaluation**, continue to refine them and add new ones.
5. **Implement** both impact and process evaluations: measure impacts and test

hypotheses related to alternative explanations based on analyses of contextual factors (biophysical, social, cultural, economic, political, institutional; i.e., strong inference; Platt 1964).

6. **Elucidate** whether the intervention caused the observed impacts (attribution and contribution analyses).
7. **Assess** threats to the validity of the evaluation.
8. **Synthesize** evidence for the impacts.
9. **Support** the use of the new knowledge gained through the evaluation.

Findings of the impact evaluation need to be built through cross-disciplinary integration of evidence corresponding to the different domains that underlie responsible forest management. As with the results of the process evaluation (Steps 5–7), the findings will represent diverse forms of evidence that need to be integrated to produce

an evaluation judgment that participants in the evaluation process will discuss further. Through the iterative processes of deliberation and discussion, synthesis documents will report findings to a wide range of audiences and communicate with those within and outside the certification intervention in a transparent way (Step 8). In the spirit of enhanced capacity for change, activities that support the use of the knowledge generated to influence forest management and certification should be set into motion early in the evaluation process (Step 9). These actions might require new institutional arrangements and policies (e.g., partnership agreements, legal frameworks for monitoring and verification, refined and adapted interventions). The cycle of adaptive management will make one full turn when new experiments on resource management, including tests of novel rules and associated implementation models, are initiated (Figure 2).

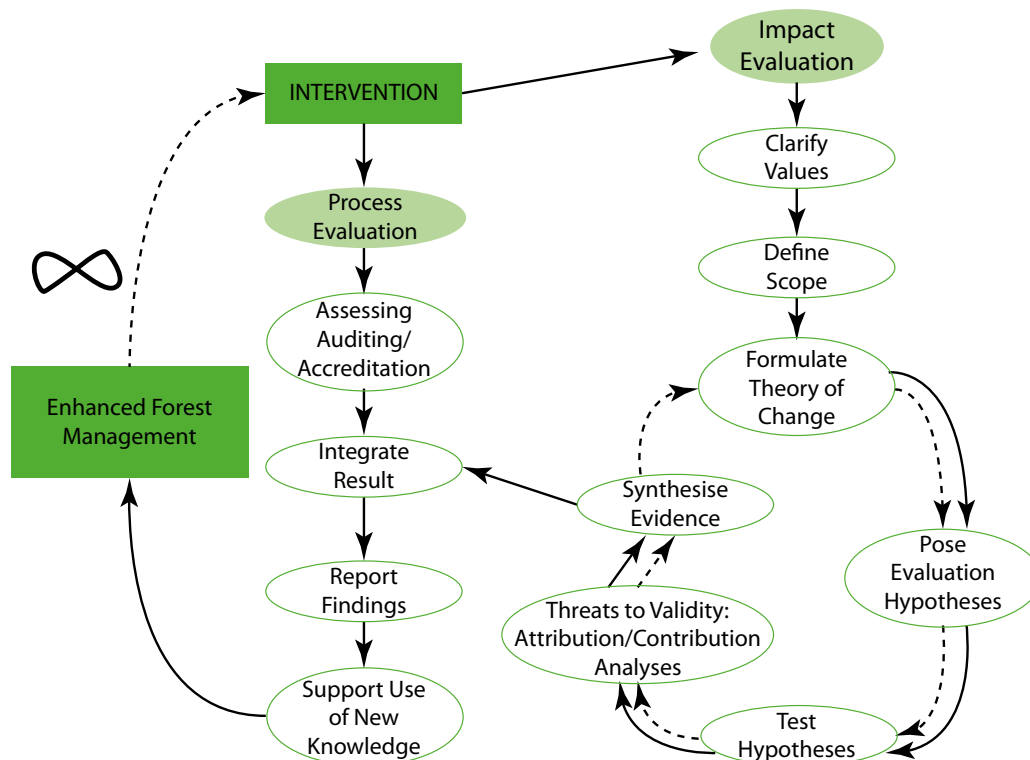


Figure 2. Complementarity of the impact and process evaluations of certification that together facilitate the generation of new knowledge in a participatory manner (i.e., social learning) and frame adaptive management of forests

Note: One result is improvements in forest management through enhanced decision-making processes and arrangements (e.g., boosted accountability, more appropriate policy arrangements and institutions, increased knowledge). Experiments in policy experimentation (e.g., adaptation and design of new interventions) for resource management and continued learning represent the foundation for adaptive management (dashed black line).

Clarify the values underpinning the evaluation of certification

Evaluation involves assessment of the impacts of an intervention based on the values and aspirations of a group of interested parties. Consequently, intended users of the results of an evaluation need to be included in the evaluation's design and then consulted regularly during its implementation.

Their involvement will also provide insights into the kinds of information needed to satisfy their needs; it might also help identify how to obtain that information. Overall, this process will generally render the evaluation more engaging, dynamic and transparent. For the evaluation of certification, at least the target audiences listed below should be included:

- **Donors** who want to know if their investments have served the intended purpose and if the theory behind certification is robust (or was more robust in the past). Donors will likely use knowledge gained as a guide to redirect their investments.
- **Government** agents who want to gauge how certification works in relation to current and future policies related to forest management (e.g., redundancy and complementarity with existing legal frameworks with similar goals such as national regulations and international efforts at legality assurance).
- **Certifying bodies** that want to learn how their work contributes to forest management sustainability and to become aware of deficiencies in their work.
- **NGOs** that support certification but recognise the need for objective and independent measures of its contributions to achievement of their own goals.
- **FMUs** whose managers are interested in learning how their operations and decisions related to certification fit into the larger picture of forest management sustainability.
- **Other stakeholders**, especially people who live near FMUs or whose welfare is otherwise affected by forestry operations, and who should benefit from the improved management of certified FMUs, as well as society at large, including consumers who purchase certified goods.

Define the scope of the evaluation

Although this paper focuses on the evaluation of natural forest management in the tropics, similar work is needed in temperate forests and for planted forests. Some of the impact questions pertain

mostly to the FMU level, whereas others need to be addressed at larger spatial scales through remote-sensing and other techniques. To define the scope or set the intellectual and geographical boundaries of the evaluation, deliberative processes involving key stakeholders should focus the evaluation efforts on the most important questions.

Impact and process evaluations

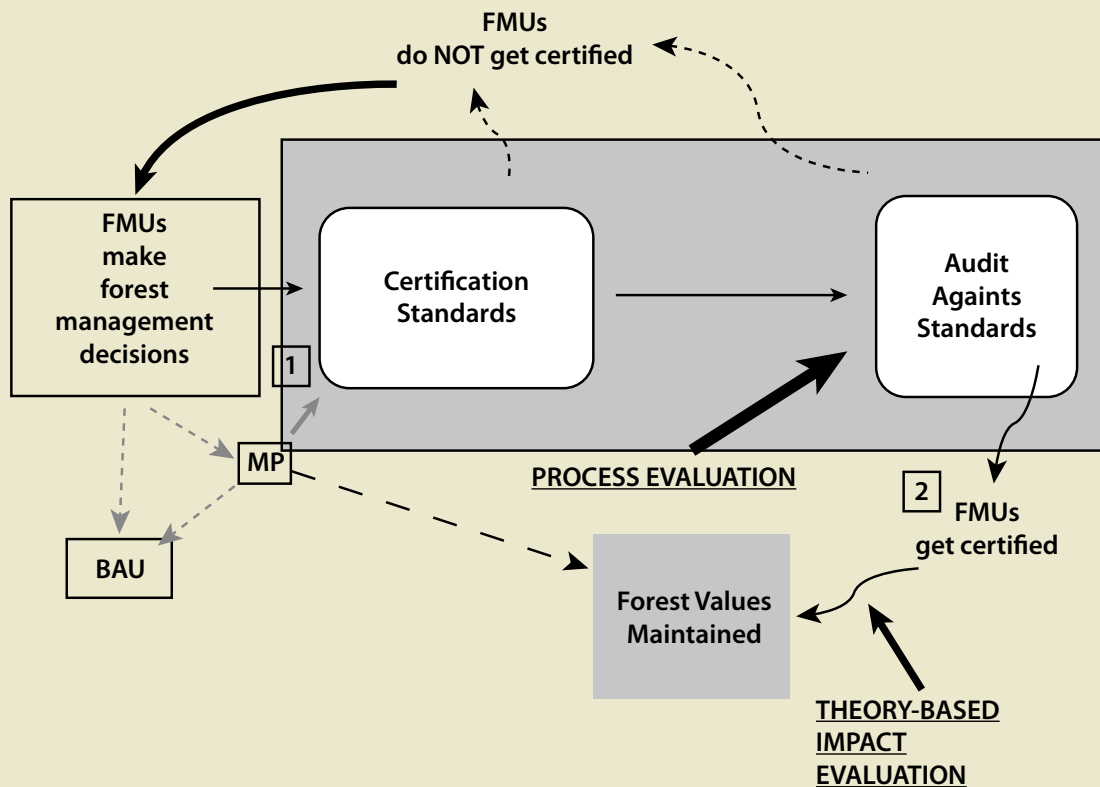
Detailed understanding of two different processes is needed to enable a proper evaluation of certification: evaluation of the theory behind the intervention through empirical examination of outcomes (theory-based impact evaluation) and process evaluation (Figure 2; Box 2). Theory-based impact evaluation explores whether impacts are achieved due to the intervention (i.e., if the required practices adopted by managers and audited by certifying bodies actually maintain forest values; SCR 2012; Stern *et al.* 2012) and to establish whether failure to achieve goals occurred even when the intervention was properly implemented. In other words, theory-based impact evaluation will reveal whether adequate compliance with certification standards secures achievement of intended outcomes. For this purpose, detailed measures of outcomes are needed (e.g., reduced deforestation or the maintenance of timber yields; see Table 3 for a non-exhaustive list; this process corresponds to the dark grey portion in Box 2). Theory-based impact evaluation would require a theory-of-change for forest certification and identification of its underlying risks and assumptions; this tool would propose impact pathways through which change is achieved on the ground.

Notes on a theory-of-change for FSC certification

Development of a theory-of-change to guide the evaluation of certification requires synthesis of existing knowledge and a *a priori* examination of the evaluation questions to be addressed. A theory-of-change is also the product of critical thinking about how and why changes occur, while clarifying expected outcomes and pathways through which change happens (Retolaza 2011; Guijt and Retolaza 2012; Stein and Valters 2012; SCR 2012). A key step towards building a theory-of-change is to assure that the process is participatory (Rogers 2012; Vogel 2012). An iterative approach is also needed so as to make as explicit as possible the intended and realized impacts of the certification intervention (James 2011).

Box 2. FSC certification and evaluation approaches

the logic behind an FMU's decision to manage rather than simply exploit a forest for timber is the first link of the model of change that can lead to FSC certification. Even if an FMU prepares a management plan (MP), it might decide to disregard its plan and carry out logging operations in the business-as-usual (BAU) manner. Alternatively, if the MP is of good quality and the FMU follows it, many of the goals of certification might be reached without a certificate being granted or even sought (dashed black line). Another possibility is that the FMU contracts a certifying body (CB) to make a pre-assessment visit but then realizes the costs of required changes would exceed the expected benefits from certification and abandons further efforts to become certified (dotted line). If the FMU does pursue certification, the next step is to implement on-the-ground actions to comply with certification standards. To become and remain certified, these improved practices are verified through the auditing process by a CB on an annual basis (*process evaluation*). After being certified, the FMU might allow its certificate to lapse or have it rejected due to failure to address CARs. FMUs that remain certified must still demonstrate their management practices lead to desired outcomes (*theory-based impact evaluation*). The FMU can then revisit the choice of forest management regime based on a suite of factors and decide to apply for certification, return to BAU management practices or maintain improved management practices.



A theory-of-change also clarifies how certification is intended to promote management practices that maintain forest values under different dynamic scenarios (e.g., unpredictable markets) and clarifies barriers to its successful implementation (e.g., financial, knowledge, technical; Furman 2005; Rugh *et al.* 2010; White 2010; Gertler *et al.* 2011; Rogers 2012). It specifies the inputs (i.e., financial, human and material resources needed

for the intervention), activities (i.e., actions to be taken or work performed through which inputs, such as funds, technical assistance and other types of resources, are mobilized to produce specific outputs), outputs (i.e., the products, capital goods and services that result from the intervention, as well as any changes resulting from the intervention that are relevant to the achievement of outcomes) and outcomes (i.e., the expected or achieved short-

Table 3. Examples of expected outcomes of certification to be examined to determine the impacts of forest management certification

	Biophysical	Social/livelihoods	Economic	Governance/policy
Outcomes	RIL practices are used.	No accidents from logging practices.	A fair distribution of benefits derived from the FMU is discussed and accepted among concerned parties.	Effective control of access and proper use of allowed resources by FMU and corresponding institutions exist.
	Fires are appropriately managed/controlled.	Negotiation processes (i.e., Free Prior and Informed Consent (FPIC), platforms and mechanisms for conflict avoidance) are effective, efficient and legitimate.	Legal compensation for the exploitation and establishment of the FMU occurs in a fair and timely manner for all.	Stakeholder (e.g. FMUs, governments, and citizens) are aware of timber and other forest values.
	Probability of deforestation is reduced.	Land and resource rights of local communities are not alienated.	FMU managers are not charged with administrative or penal sanctions.	Explicit, transparent and legitimate ways to negotiate with stakeholders exist.
	Spatial continuity of forest cover is maintained (corridors).	Workers have job security and access to legal services.	FMU managers seek to improve operational efficiency/minimize harvesting waste.	Clear sanctions and appropriate penalties are in place.
		Workers aware of sustainable forest management needs/goals.	Taxes and royalties are paid in a timely manner.	There is compliance of legal frameworks and procedures to avoid corruption.
		Workers' housing meets minimum national standards.		Required taxes, fees and royalties are paid in a timely manner.
		There is fair benefit-sharing with local communities.		
Outputs	Hydrological functions and services (e.g., flow regimes and water quality) are maintained.	Worker welfare is enhanced.	FMU is financially viable to its owner (e.g., government).	Legal/institutional frameworks are effective, efficient and equitable.
	Biodiversity is maintained at the genetic, population and landscape levels.	Livelihoods of surrounding communities are enhanced.	Operations in the FMUs guarantee the compatibility of forest resource management with ecosystem service provision.	There is upward accountability ^a .
	Productive capacity is not impaired and future harvest volumes are secured.		The values of timber stands are maintained.	There is downward accountability ^b .
	Carbon emissions/pollution minimized.		FMU is financially viable to its owner.	
	Ecological processes are not threatened by forest management. HCVFs, buffer zones and riparian set- asides provide biodiversity and other benefits.			

a To reassure product consumers, donors, taxpayers, decision-makers and society at large that resources are being wisely used and/or invested (Rogers 2012).

b To inform knowledge users (intended beneficiaries and communities) about whether or not and in what ways a program benefits the community (Rogers 2012).

Note: Shaded boxes of the same color represent sets of outputs and outcomes that can inform on more than one dimension of certification (e.g., gateway outcomes; A. Rowe pers. comm.). RIL= Reduced-impact logging. HCVF = high conservation value forest.

term and medium-term effects of an intervention’s outputs) towards the goal of responsible forest management (all definitions according to OECD 2002 and Jagger *et al.* 2010).

A well-crafted theory-of-change may also help clarify which aspects of the intervention are needed and which behavioral and contextual variables need to be addressed to explain variation in the intervention impacts (i.e., impact pathways). In the context of each study site, the theory-of-change helps to reveal the particular roles and interactions among contextual factors (e.g., time of first implementation of certification; existing legal frameworks; capacity and technical knowledge of the staff; timber market characteristics) and the associated actions related to the certification intervention (e.g., involvement of local stakeholders; improvements in timber harvesting practices). The resulting models of change, tailored to particular regions and locales, should reflect specific issues that need to be tracked to assess the impacts of certification. To date, none of the

existing natural product certification systems has an explicit and available theory-of-change that describes the change process, but Fairtrade recently started working on one (Nelson and Martin 2011).

An example related to the implementation of measures within FSC’s Criterion 4.1 (‘Forest management should meet or exceed all applicable laws and/or regulations covering health and safety of employees and their families.’ FSC 2012) is used here to guide the elaboration of a theory-of-change (Figure 3). In this example, there are various sources of motivation for FMUs to reduce work-related injuries, including: compliance with national regulations (purple color; Figure 3); pressure from worker unions, trade associations or stockholders; public awareness; concerns about corporate reputation; genuine concern about worker welfare; and, the need to comply with FSC criteria. A proper impact evaluation design will help pose research hypotheses to tease out the effect of these and other factors.

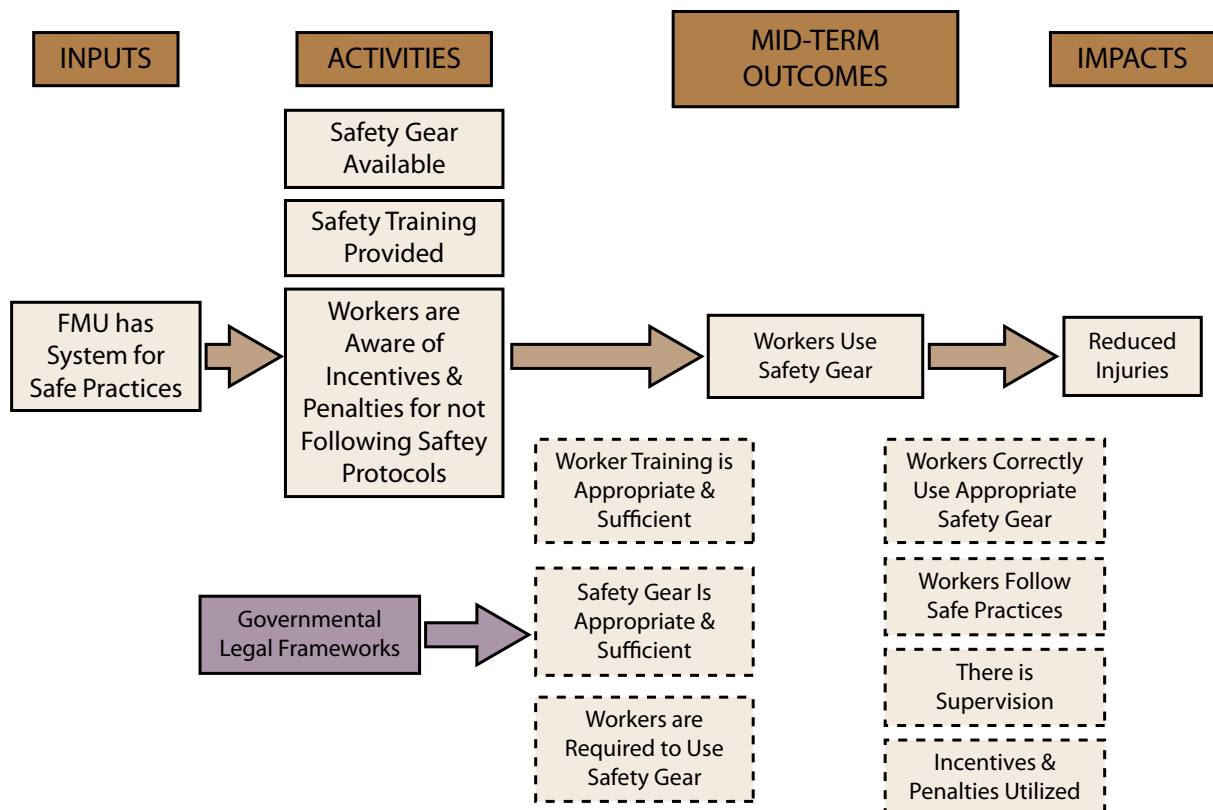


Figure 3. Example of a specific theory-of-change that accounts for reduction of work-related injuries in a certified FMU

Note: Boxes with dashed outlines represent assumptions behind the transition from one step to the next in the sequence of actions that lead towards the long-term goal of reduced work-related injuries. Dashed boxes could also be channels for certification impacts. For example, the law may nominally require safety plans and safety equipment (purple box and arrow), but it is certification that ultimately encourages the FMU to require/provide incentives for workers to use the equipment.

Table 4. Risks and assumptions that underlie the transitions between outputs into short-, medium- and long-term outcomes from the forest certification intervention

	Outputs to short-term outcomes	Short- to medium-term outcomes	Mid-term outcomes to impacts
Risks	<ul style="list-style-type: none"> Market signals fail to drive demand for certified products. Lack of financial support to implement practices required for certification. Workers refuse to wear safety gear. 	<ul style="list-style-type: none"> Communication with local communities and institutions is insufficient and conflicts abound. Implementation of certification is too complicated. 	<ul style="list-style-type: none"> Certification does not lead to responsible forest management. Impossible to sort out tradeoffs between competing components of forest management sustainability. Lack of continuous financial support undermines the intervention.
Assumptions	<ul style="list-style-type: none"> Subsidizing certification (e.g., funds for training and planning) leads to full adoption of the scheme. Training translates into better implementation of recommended forest management practices. Full information on available resources improves forest management decision-making. 	<ul style="list-style-type: none"> FMU controls access to protected areas, HCVFs, and other set-asides. FMU internalizes good forest management practices. Appropriate social contracts are defined between FMUs and local institutions. 	<ul style="list-style-type: none"> Financial benefits (direct and indirect) of certification exist. Forest management certification is integrated with other policy instruments aimed to improve forest management. Green markets are created and FMUs obtain incentives that more than compensate for their certification-related efforts.

Note: Subsequent changes in impacts will depend on how risks and assumptions are considered during implementation of the certification intervention.

FMU = Forest Management Unit
 HCVF = high conservation value forest

In parallel with formulation of a theory-of-change for certification, a participatory analysis of *risks* and *assumptions* is needed (i.e., necessary conditions for change; Wigboldus and Brouwers 2011; Stein and Valters 2012). This analysis can reveal factors that might hinder the transition from *Outputs* to desired *Outcomes*. Also, assumptions about how the biophysical, social, economic and policy change processes occur, as well as specification of the impact pathways, all need to be explicit and addressed throughout the evaluation. Some apparent risks and assumptions for FSC certification are presented in Table 4.

Process evaluation

Process evaluations (i.e., implementation evaluations; Weiss 1997; Rossi *et al.* 2004; SCR 2012) seek to determine whether all specified steps for certification were implemented as

specified by the intervention. They also seek to ascertain if the intervention was implemented differently in different places (Rogers 2012). Steps in the certification process include preparation of a management plan as specified by the FSC certification standards; implementation of practices in accordance with the management plan; and, most critically, verification of compliance with the standards through an audit (this process corresponds to the light grey portion in Box 2). For a proper process evaluation, it is thus critical to know a great deal about the company, the auditors, auditing processes and certifying bodies.

Certification audits include desk and field assessments, as well as consultations with stakeholders (e.g., public consultations). Evaluators of forest certification will need to understand how certification guidelines are interpreted and

implemented in the field by the auditors and how auditor performance is verified by the certifying bodies. They will also need to be familiar with how FSC conducts its annual audits of certification bodies, a process that includes random checks of documentation as well as field verifications. Given the critical importance of the judgments of auditors to the success of certification, more needs to be known about who they are, the conditions and constraints under which they work, and how they make decisions (Cashore and Auld 2012).

Accreditation Services International (ASI, established in 2006; www.accreditation-services.com), which provides third-party accreditation of certification bodies, represents an additional layer of verification of compliance with FSC guidelines. Before 2006, the FSC accredited directly its certifying bodies. Currently, ASI determines the competence of certifying bodies through the review of secondary information, periodic visits to the offices of the certifying body and witness assessments (observations of auditors by ASI personnel). Accreditation of a certifying body involves an initial assessment, annual surveillance to verify implementation of recommendations, and visits to certificate holders by ASI representatives to verify compliance (i.e., an audit of the audits; ASI 2010, 2011).

Evaluation questions

Generally, evaluations address high-level impact questions (e.g., did certification deliver its intended

impacts?), as well as questions more centered on the parties affected by the intervention (e.g., for whom and under what conditions did certification work?). They also consider both positive and negative outcomes. More scaled-down questions address types of impacts and their distributions. For example, it is important to know how the impacts of certification have changed over time. Also needed is information about the influence of other factors on certification impacts (e.g., how did certification interact with other initiatives to achieve or fail to achieve desired outcomes?). More fundamentally, it needs to be clear how, why and where the intervention worked, whether there was variation in implementation of the certification process, and to what extent any differences in impacts can be explained by this variation (adapted from Rogers 2012; Tøn *et al.* 2012).

Further grounded questions that might be linked to particular places of interest are how certification schemes and certifying bodies compete among themselves and how such competition affects the performance of the FSC certification (e.g., to avoid having to make required changes, some FMUs might prefer to switch to CBs within FSC or switch from FSC to an alternative certification system). The specific questions to be asked through the empirical evaluations should emerge from consultations with a range of relevant stakeholders from each site, and might represent a refined list of those mentioned above.

4 Knowledge needs for the design of an evaluation of the impacts of forest management certification

This section describes information needed to inform discussions and negotiations related to the design of the evaluation of FSC certification impacts. Specific analyses of certain themes may shed light on critical components to inform that design; detailed information on these issues is presented. Results of these analyses are meant to provide a foundation for carrying out the evaluation.

Knowledge about the forest sector

An evaluation of the impacts of forest certification should take into account that, even in the complete absence of this intervention, FMUs vary in the quality of their management. For example, some may employ practices that closely match certification requirements (e.g., employ reduced-impact logging techniques and collaborate with neighboring communities). As a further complication, these practices vary over time in response to a wide variety of factors of which certification is only one (e.g., extent of enforcement of forest regulations, timber market dynamics). It should also be noted that the process of certification involves many steps over several years (e.g., Ruslandi *et al.* in press) along with many possible shortcuts, particularly if the FMU secures support from NGOs. It is also important to note that both getting and remaining certified requires substantial investments of time, money and human resources. Currently certified FMUs vary broadly in when they were first certified as well as in the numbers and types of adjustments in management practices required to obtain or retain their certificates. For example, a study of five certified forest concessions in Kalimantan, Indonesia, revealed that it took from 3 to 10 years to achieve certification (Ruslandi *et al.* in press).

Variation in management practices and certification status can be portrayed as a continuum – from those FMUs that employ sub-standard management (*sensu latu*) practices and have no interest in certification to those already certified

through several rounds of audits (Figure 4). Some temporal dynamics in management practices and certification status (e.g., certification lost, certification pending) are also possible and are thus represented along this continuum (based on previous research of forest sector evolution and legal regimes in the tropics; Ruiz-Pérez *et al.* 2005; Salazar and Gretzinger 2005; Pereira *et al.* 2010). Of particular relevance are changes in macro-economic conditions, currency exchange rates and land tenure arrangements (Cattaneo 2001; Wunder 2005; Arcand *et al.* 2008; Sunderlin *et al.* 2008; Banerjee and Alavalapati 2009). A typology that captures where FMUs are on this continuum is needed to inform the design of empirical evaluation of impacts. Such a typology will also provide information needed to assist in the selection of FMUs as counterfactuals and will generate insights about factors that might directly or indirectly influence the impacts of certification. To this end, up-to-date information about the pool of existing FMUs and the contexts in which they operate needs to be compiled. The variables used to generate the typology are based on previous efforts at characterizing forest sectors (Table 5).

Knowledge about the temporal dynamics of certification

The position of an FMU along the certification continuum (Figure 4) can change over time in response to political economy factors related to the timber and other sectors, investments, opportunities, market realities and other drivers. In particular, contextual factors that operate at local, national and international levels can influence an FMU's decisions to opt for certification and, once certified, to remain so. Market dynamics (e.g., consumer preferences and acquisition power) change and, in turn, influence suppliers' decisions *vis à vis* certification. Shifting legal frameworks and their enforcement, technical capacities, technological innovations, global/regional/national economic conditions and cost–benefit ratios are among the factors that can affect FMU decisions

Table 5. Variables that affect forest management

Type of characteristic	Attribute
Exogenous to the Firm	
Economic	Origin of firm (country)
	Origin of capital (country)
	Vertically integrated (firm owns milling and other processing facilities: YES, NO)
Political	Institutional regime (e.g., community, private, public)
	Legal framework: evolution and dynamics
Exogenous to the FMU	
Biophysical	Area (ha)
	% area with slopes > 10, 20 and 30°
	Previously logged (YES, NO)
	Distance or travel time to harbor/mill
Social/Livelihoods	Population density in surrounding area (#/km ²)
	Dominant ethnic group(s) in area
	Recognized resource use and tenure rights of local communities (YES, NO)
Political	Existing conflicts with communities or other stakeholders (YES, NO)
	Administration regime (e.g., district, state)
	Type and duration of harvest permit
	Political cohesion of industry sector
Endogenous to the FMU	
Biophysical	Area logged/yr. (ha/yr.)
	Volume harvested/yr. (m ³ /yr.)
	Logging intensity (range and mean; m ³ /ha)
	# species marketed
Social/Livelihoods	Workers: origin, gender (#, %)
	Potential conflicts with FMU, within communities, with local stakeholders (YES, NO)
Economic	Market orientation (principal outlet: national/regional, Europe, North America, Asia)
	Logging subcontracted (YES, NO)
Political	Management status (certificate of legality, approved management plan, RIL-certified, no official management status)

Note: The table illustrates examples of variables likely to influence the short- and long-term outcomes of forest management and thus the impacts of certification. Data collected with this framework will be used to construct typologies of FMUs and thus to inform the dynamics of certification and the process of self-selection into the scheme.

FMU = Forest Management Unit

RIL = Reduced-Impact Logging

about remaining certified or not (Nebel *et al.* 2005; Kollert and Lagan 2007; Crow and Danks 2010; Chen *et al.* 2011b). Substantial understanding of these dynamics will reveal the true impacts of forest management certification. Proposed categories are *Never Certified*, *Considering Certification*, *Enroute to Certification*, *Certified*, *Certification Failed* and *Certification Lost* (Figure 5). By *Never Certified*

we refer to FMUs with no apparent interest in certification and for which certification, under current contextual conditions, might represent too much of an investment risk or a waste of resources.

Understanding the contextual factors that influence forest management decisions in general and certification in particular should help explain

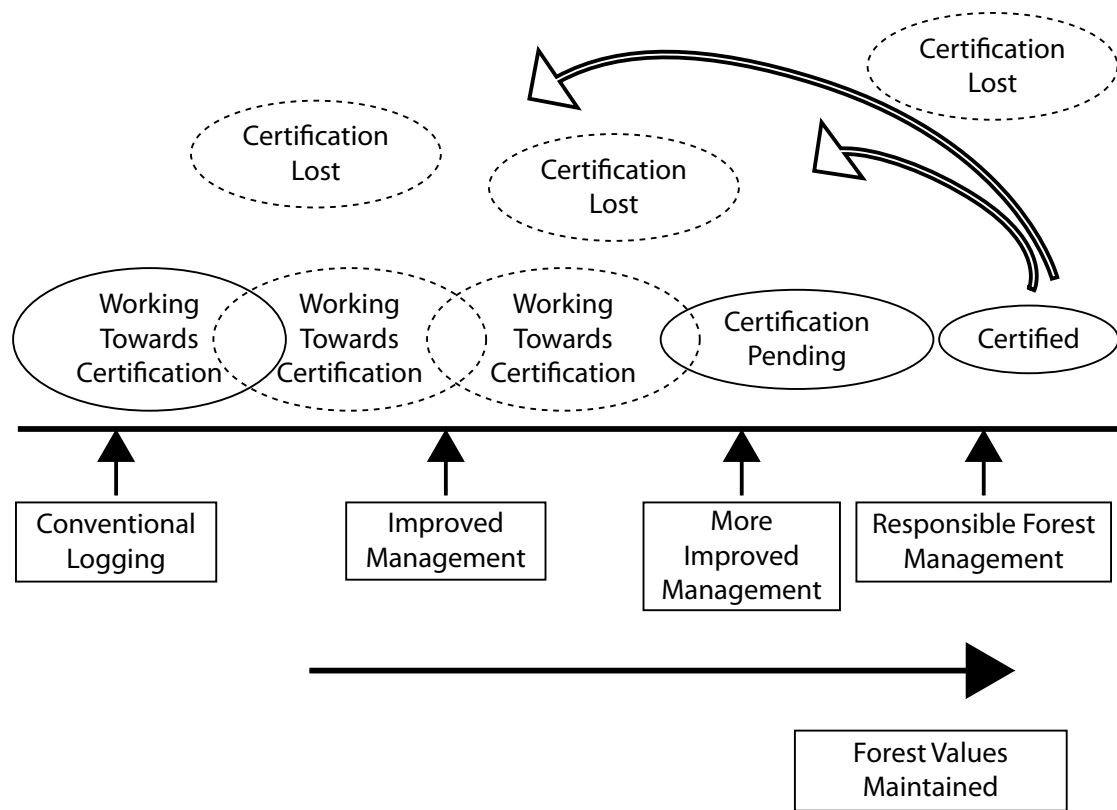


Figure 4. The certification continuum

Note: The certification continuum superimposed on an axis of responsible forest management reflects different stages related to certification (ovals). An FMU that loses its certificate might end up at different places along the axis depending on whether it continues to employ the practices required for certification. It is worth noting the possibility that FMUs not interested in certification employ responsible management practices that lead to the maintenance of the biophysical, social, economic and political values of forests (Box 2). Also, FMUs working towards certification employ a range of different quality practices (ghosted ovals).

the observed transitions in the certification status of FMUs. It might help to focus on the most common transitions and their temporal patterns to reveal related contextual elements (e.g., changes in legal frameworks that influenced FMU choices, information availability, market crashes). Any possible switches among certification schemes (e.g., FSC to PEFC) or certifying bodies (e.g., SGS to Control Union) and the effects of these switches on forest management outcomes should also be explored and explained.

Knowledge about the drivers of self-selection

Knowing why FMUs opt for certification – including what factors affect the probability of obtaining certification – is critical to the evaluation process. Factors that influence the self-selection process (i.e., facilitating factors and barriers); their

domain (i.e., internal such as costs, expertise and knowledge vs. external such as consumer confusion and supply-chain structure); and their dynamics should also reveal much about FMU transitions along the certification continuum in Figure 4. In-depth understanding of FMU decision-making processes about certification will reveal the relative roles of factors such as legal contexts, technology access and information availability. Diverse factors influence a FMU's decision to seek and retain certification; they differ in their effects among countries, regions and types of FMUs, and vary over time. FMUs make decisions about certification based on the economic and other sorts of costs and benefits (Richards 2004). Key inputs for this assessment are the results of the typology and data on the dynamics of the forest sector (discussed in the previous sections), which will provide opportunities for comparing information obtained with contextual factors. A recent report

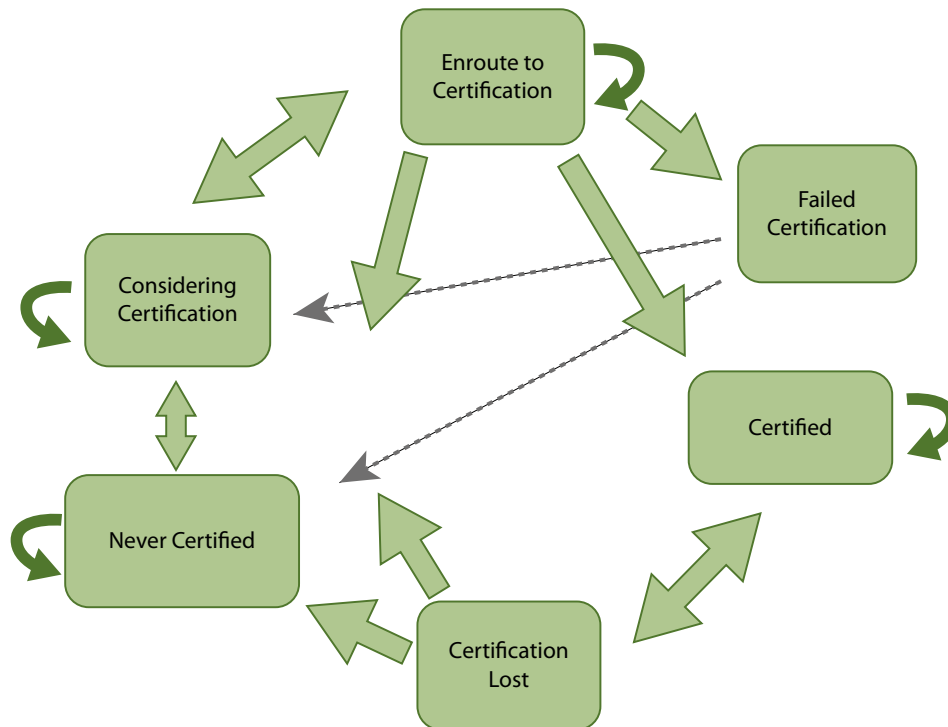


Figure 5. Contextual and other factors that influence FMUs' management decisions

Note: Contextual and other factors (e.g., retention of trained workers; availability of capital) influence the choices FMUs make about certification over time. Arrows represent the transition probabilities (or proportions) of FMUs that remain in each category (curved) and that move into other categories (straight including dashed lines) during a given period of time.

that proposes a typology of FMU attitudes towards certification (SCR 2012) offers valuable reflections on self-selection. These attitudes, in combination with other factors (e.g., estimated cost–benefit ratios) influence FMU decisions regarding certification. This framework is used to outline current knowledge about self-selection into certification (Table 6).

Several studies have attempted to assess the direct and indirect costs of certification as a potential barrier to participation. Some direct costs vary with baseline FMU management practices. For example, FMUs that employ good management practices will need to invest relatively little to satisfy certification standards. But even in relatively well-managed FMUs there are often deficits in the knowledge, skills and practices needed to reach certification standards. Apart from improvements in management practices and record-keeping, other costs include equipment needs, staff training and additional salaries, audit and membership fees, monitoring and record-keeping, and consultation processes, some of which vary with FMU size (Cashore *et al.* 2006; Gale 2006; Roberts 2012).

The financial costs of certification vary substantially. In a study of five FSC-certified concessions in Indonesia, for example, Ruslandi *et al.* (in press) estimated costs of certification of USD 300 000 – 700 000 per concession with averages of USD 4.76/ ha and USD 0.1/m³, half of which was covered by outside agencies. In general, these costs are lowest for FMUs that already employ sound management practices, which is why they may be among the first to seek certification (Blackman and Rivera 2011; Blackman and Guerrero 2012). Likewise, due to economies of scale, the ability to cover direct and indirect costs (and thus effectively join certification schemes) increases with the size of the FMU (Nussbaum *et al.* 2001; Rivera 2002; Roberts 2012). In recognition of this bias against small FMUs, many of which are community run, the FSC started the Small- and Low-Intensity Managed Forest (SLIMF) program in which FMUs pay reduced rates for initial and follow-up audits (e.g., Brazilian FSC SLIMF standard, FSC 2013).

Possible indirect costs of certification not considered above derive from foregone harvest

Table 6. Typology of attitudes regarding forest management

Attitude	Reasons	Rationales	References
Tactical risk management	<ul style="list-style-type: none"> - Perceived threats of more stringent governmental regulations. - Becoming certified might facilitate further dealings with the government. - Certification makes FMUs less prone to scrutiny. - Increase acceptance of FMUs by stakeholders. 	<ul style="list-style-type: none"> - Especially important if the FMU has been sanctioned in the past. - Some FMUs reportedly avoid certification precisely to evade attention (E. Meijaard, pers. comm.). 	<ul style="list-style-type: none"> Rivera 2002 Rivera and de León 2004 Blackman and Guerrero 2012
Tactical opportunity	<ul style="list-style-type: none"> - Secure and enhance market access; importance varies with FMU size and market orientation. - Expectation of price premiums. - Expectation of other benefits (e.g., reduced transaction costs). 	<ul style="list-style-type: none"> - FMUs anticipate economic benefits and once certified, can sometimes capitalize on them. - Some benefits are widely realized and others (e.g., price premiums) are restricted to specific conditions. 	<ul style="list-style-type: none"> Rivera 2002 Hartsfield and Ostermeier 2003 Nebel <i>et al.</i> 2005 van Kooten <i>et al.</i> 2005 Shahwahid 2006 Kollert and Lagan 2007 Crow and Danks 2010 Chen <i>et al.</i> 2011b Prokopovych 2012
Strategic risk management	<ul style="list-style-type: none"> - Use certification to buffer market risks. 	<ul style="list-style-type: none"> - Reduced operating costs due to improved forest management operations and better trained personnel. - Enhanced learning and transparency. 	<ul style="list-style-type: none"> Takahashi <i>et al.</i> 2003 Pokorny and Steinbrenner 2005 Auld <i>et al.</i> 2008 Vidal and Kozak 2008 Araujo <i>et al.</i> 2009 Cubbage <i>et al.</i> 2010
Strategic opportunity	<ul style="list-style-type: none"> - Certification as an avenue to more responsible environmental, social and economic productive systems. 	<ul style="list-style-type: none"> - Increased public confidence and social acceptance. - Greater environmental and social responsibility associated with certification. 	<ul style="list-style-type: none"> Cubbage <i>et al.</i> 2010

Note: A combination of these and other factors underpin the impacts of forest management.

FMU = Forest Management Unit

volumes due to restrictions on logging in riparian buffer zones and other set-asides; increased durations of harvest cycles; restrictions on harvesting during wet weather; and requirements for seed tree retention and other felling restrictions to ensure natural regeneration and maintain wildlife habitat (Cubbage *et al.* 2003; Simula *et*

al. 2004; Gan 2005; Chen *et al.* 2010). Indirect and often hidden costs of dealing with governance failures (e.g., conflict over resource access and paying bribes) and the costs of negotiations with local and regional stakeholders are generally also hard to capture in considerations of the costs of certification.

Due to the scarcity and modest size of price premiums on certified timber (Cashore *et al.* 2004; Fischer *et al.* 2005; Gale 2006; Cerutti *et al.* 2012), the factors motivating FMU participation in certification are often not directly financial in nature. For those five concessions in Indonesia mentioned above, Ruslandi *et al.* (In press) suggest the following motivations for certification: concerns over corporate reputations; interest in improving operational efficiency through better forest management practices; hopes that certification will benefit FMU governmental and non-governmental relations; and, continued hopes for improved market access and price premiums. Clearly, a complete analysis of FMU self-selection into certification would need to consider the external financial and technical support provided to FMUs of all sizes and types. In the Ruslandi *et al.* (In press) study, for example, organizations that funded some direct and indirect costs of certification include NGOs (e.g., World Wildlife Fund and The Nature Conservancy), development assistance agencies of several governments (e.g., The Borneo Initiative from the Netherlands) and research and training institutions (e.g., Tropical Forest Foundation).

Knowledge about the political economy of the forestry sector

A final issue related to the evaluation of certification pertains to the contexts in which FMUs are embedded. Certification is applied in a diversity of places that vary in characteristics that can affect the outcomes of the intervention and how they change over time. Examples include governance issues such as changes in legal frameworks and their implementation, as well as changes in the policy environment (e.g., timing of elections, changes of government); social issues related to changes in capabilities and extents of participation; economic factors that vary with market structures and preferences, supply chain structure and dynamics, technology and availability of capital; and biophysical aspects of forest management that range in resource availability as affected by harvesting history, intensity, disease, fire and climate change.

Conversely, it is also important to consider the direct and indirect impacts of the certification intervention in the context in which it is applied. For example, certification principles inspired

national forest policy-making in Bolivia (Nittler and Nash 1999), while forest certification contributed to enhanced forest law compliance in Cameroon (Cerutti *et al.* 2011). In Russia, certification helped to build adaptive capacity by increasing knowledge of local stakeholders through more participative forest management processes (Keskitalo *et al.* 2009). Thus, a solid grasp on how policies and programmes evolve and shape decisions regarding forest management is needed; this can help to adequately contextualize the outcomes of forest management certification. Proper understanding of the political economies in which forest management occurs may also aid the design of robust evaluations of certification.

Knowledge about land use change

Blackman (2012) and Gaveau *et al.* (2012) point out that remote-sensing techniques can help measure how certification reduces deforestation at a diversity of scales. For example, certification might affect the probability of deforestation at the FMU level, at the level of a company that manages several FMUs or at the jurisdictional level where it is located; this is one of the objectives of certification (Viana *et al.* 1996). Availability of cloud-free images remains a challenge in some parts of the tropics, but the approach is well-established and regional expertise at least is generally available. The impacts of construction of logging roads, opening of skid trails and even felling can also be monitored remotely. However, the required techniques (e.g., ClasLite: www.claslite.stanford.edu) are sophisticated and will not be considered further here.

Remote sensing can reveal changes in forest cover in FMUs along the certification continuum, but a theory-of-change is needed to elucidate causes for observed changes. For example, lack of governmental support for enforcement of FMU rights or high costs may constrain FMUs from stopping deforestation (Coleman and Steed 2009; McElwee 2010; Amacher *et al.* 2012). Conversely, a remote-sensing study might miss a more insidious cause of forest loss: when boundaries of FMUs are redrawn to allow portions to be cleared legally. Such an impact will be revealed only if both the original and final FMU boundaries are known. Remote sensing will also not reveal the causes of deforestation within FMUs that retained their certificates. Finally, remote-sensing studies

of the effects of certification on deforestation still need appropriate counterfactuals; these would determine whether or not deforestation would have happened in the absence of certification. As such, to infer the causal effects of certification on observed land use changes, remote-sensing evidence needs to be interpreted with reference to the full set of contextual and site-specific characteristics used in the forest typology and the construction of a theory-of-change.

Previous efforts at assessing certification impacts on deforestation used panel data from the Food and Agriculture Organization of the United Nations (FAO) for 1972–1994 and 2005–2010

for developed and developing countries (Damette and Delacote 2011). Potentially explanatory variables included in the analyses were the volumes and values of harvested timber. Control variables included national-level indicators of institutional quality (i.e., indices of political rights and civil liberties); a model of deforestation as a function of the country's GDP; annual GDP growth rate; population density; and the country's forest cover at the beginning of the study period. Results from this study suggest that countries in which FSC certification was prominent experienced less deforestation (Damette and Delacote 2011), but did not address the causal mechanisms that could explain this outcome.

5 Concluding remarks

Forest certification is a private, voluntary, market-driven instrument designed to promote responsible forest management. While many certification systems operate around the world, this paper focused on certification of natural forest management in the tropics by the Forest Stewardship Council (FSC). The FSC certified its first tropical forest in 1994 and has certified more than 100 other natural forest management units (FMUs) in the tropics since then. However, the often-claimed environmental and social benefits of certification remain to be empirically evaluated. After reviewing the literature on the impacts of certification, a foundation is laid here to develop an evaluation approach of what many observers claim as the most effective tropical forest conservation intervention ever initiated. The next stage is to critically analyse and synthesize information from studies outlined in the previous section and compile lessons learned and knowledge gathered into a design for an evaluation.

To be credible, salient and effective, an evaluation of the impacts of certification needs to be participatory from the onset. Input from the full gamut of relevant stakeholders coupled with compilation of the salient biophysical and socio-economic characteristics of certified and other FMUs will form the basis for designing various formal methods of evaluation of this complex intervention. Among these activities, emphasis should be given to formulating a series of theories-of-change for the intervention that captures contrasting views and angles to certification.

The evaluation framework proposed here aims at tracking variation in the quality of implementation of certification (i.e., process

evaluation). Likewise, it integrates results of empirical research that tests hypotheses motivated by how specific contextual factors shape certification outcomes. For example, large concessions, private landowners and communities are among the FMUs that have opted for certification. This basic information organized in a typology of FMUs that will serve to produce a better understanding of dynamics in the certification sector and the associated analysis of the self-selection process of FMUs into and out of certification (or switches among certification schemes/bodies). In-depth, historical political economy appraisals of the forest and timber sectors help to analyse contextual factors and other elements exogenous to the firms that manage FMUs and to the FMUs themselves. The resulting studies are meant to explain behaviors observed in the certification dynamics and self-selection analyses, to formulate hypotheses to guide the evaluation, and thus to justify the evaluation design.

Finally, a diversity of stakeholders – from representatives of local and regional communities and governments, environmental and social NGOs to FMUs at all levels of decision-making – should be invited to contribute to, and to benefit from, the knowledge gained from an evaluation of forest management certification. After all, the framework proposed here is only the first step towards the design and future implementation of evaluation research in the context of tropical forest certification on a global basis. Ultimately, it is hoped the research framework proposed will contribute to learning from past mistakes, building on lessons learned and enhancing decision-making towards the maintenance of forest values over the long term and for the benefit of society as a whole.

6 References

- Accreditation Services International (ASI) 2010 Guidance on witness assessments. ASI-GUI-20-105. Witnessing v 3.0. Bonn, Germany.
- Accreditation Services International (ASI) 2011 Service delivery report. Top down case studies for ASI/FSC. Bonn, Germany.
- Agrawal, A. 2007 Forests, governance and sustainability: Common property theory and its contributions. *International Journal of the Commons* 1: 51–76.
- Agrawal, A., Chhatre, A. and Hardin, R.D. 2008 Changing governance of the world's forests. *Science* 320: 1460–1462.
- Alix-Garcia, J.M., Shapiro, E.N. and Sims, K.R.E. 2012 Forest conservation and slippage: evidence from Mexico's national payments for ecosystem services program. *Land Economics* 88: 613–638.
- Amacher, G.S., Ollikainen, M. and Koskela, E. 2012 Corruption and forest concessions. *Journal of Environmental Economics and Management* 63: 92–104.
- Amacher, G.S., Merry, F.D. and Bowman, M.S. 2009 Smallholder timber sale decisions on the Amazon frontier. *Ecological Economics* 68: 1787–1796.
- Andam, K.S., Ferraro, P.J., Pfaff, A., Sanchez-Azofeifa, G.A. and Robalino, J.A. 2008 Measuring the effectiveness of protected area networks in reducing deforestation. *Proceedings of the National Academy of Sciences of the United States of America* 105: 16089–16094.
- Araujo, M., Kant, S. and Couto, L. 2009 Why Brazilian companies are certifying their forests? *Forest Policy and Economics* 11: 579–585.
- Arcand, J.-L., Guillaumont, P. and Jeanneney-Guillaumont, S. 2008 Deforestation and the real exchange rate. *Journal of Development Economics* 86: 242–262.
- Assunção, J., Gandour, C.C. and Rocha, R. 2012 Deforestation slowdown in the legal Amazon: prices or policies? *Climate Policy Initiative/PUC-Rio de Janeiro, Brazil*.
- Arriagada, R.A., Ferraro, P.J., Sills, E.O., Pattanayak, S.K. and Cordero-Sancho, S. 2012 Do payments for environmental services affect forest cover? A farm-level evaluation from Costa Rica. *Land Economics* 88: 382–399.
- Auld, G., Gulbrandsen, L.H. and McDermott, C.L. 2008 Certification schemes and the impacts on forests and forestry. *Annual Review of Environment and Resources* 33: 187–211.
- Bamberger, M. 2012 Introduction to basic methods in impact evaluation. *Impact Evaluation Notes # 3*. Rockefeller Foundation, New York.
- Bamberger, M., Carden, F. and Rugh, J. 2009 Alternatives to the conventional counterfactual. *Real World Evaluation: Summary of Session 713 Think Tank*. American Evaluation Association, Orlando <http://www.realworldevaluation.org>
- Bamberger, M. and White, H. 2007 Using strong evaluation designs in developing countries: experiences and challenges. *Journal of MultiDisciplinary Evaluation* 4: 58–73.
- Banerjee, O. and Alavalapati, J. 2009 A computable general equilibrium analysis of forest concessions in Brazil. *Forest Policy and Economics* 11: 244–252.
- Barbosa de Lima, A., Novaes Keppe, A.L., Maule, F.E., Sparovek, G., Corrêa Alves, M. and Maule, R.F. 2009 Does certification make a difference? Impact assessment study on FSC/SAN certification in Brazil. www.imafflora.org/arquivos/Does_certification_make_a_difference.pdf
- Bartley, T. 2007 How certification matters: examining mechanisms of influence. Prepared for a working group on the assessment of sustainability standards, certification and labels, September. Indiana University, Bloomington, US.
- Bernstein, S. 2004. Legitimacy in global environmental governance. *Journal of International Law & International Relations* 1: 139–166.

- Blackman, A. 2012 Ex post evaluation of forest conservation policies using remote sensing data. An introduction and practical guide. Resources for the Future, Discussion Paper 12-05.
- Blackman, A. and Guerrero, S. 2012 What drives voluntary eco-certification in Mexico? *Journal of Comparative Economics* 40: 256–268.
- Blackman, A. and Rivera, J. 2011 Producer-level benefits of sustainability certification *Conservation Biology* 25: 1176–1185.
- Blackman, A. and Rivera, J. 2010 Environmental certification and the Global Environment Facility: STAP advisory document. Prepared on behalf of the Scientific and Technical Advisory Panel (STAP) of the Global Environment Facility (GEF). Washington, DC.
- Bowler, D., Buyung-Ali, L., Healey, J.R., Jones, J.P.G., Knight, T. and Pullin, A.S. 2010 The evidence base for community forest management as a mechanism for supplying global environmental benefits and improving local welfare. Scientific and Technical Advisory Panel (STAP) of the Global Environmental Facility (GEF). Washington, DC.
- Bradlow, E.T. 1998 Encouragement designs: an approach to self-selected samples in an experimental design. *Marketing Letters* 9: 383–391.
- Brooks, J. S., Waylen, K.A. and Mulder, M.B. 2012 How national context, project design and local community characteristics influence success in community-based conservation projects. *Proceedings of the National Academy of Sciences of the United States of America* 109: 21265–21270.
- Burgess, R., Hansen, M. Olken, B.A., Potapov, P. and Sieber, S. 2012 The political economy of deforestation in the tropics. *Quarterly Journal of Economics* 127: 1707–1754.
- Butterfield, R., Hansen, E., Fletcher, R. and Nikinmaa, H. 2005 Forest certification and small forest enterprises: key trends and impacts, benefits and barriers. Forest Trends, Washington, DC.
- Carlsen, K., Pilegaard, H.C. and Lund, J.F. 2012 Factors affecting certification uptake –perspectives from the timber industry in Ghana. *Forest Policy and Economics* 25: 83–92.
- Cashore, B. 2002 Legitimacy and the privatization of environmental governance: how non-state market-driven (NSMD) governance systems gain rule-making authority. *Governance* 15: 503–529.
- Cashore, B. and Auld, G. 2012 Forestry Review (Appendix F). *In: Steering committee of the state-of-knowledge assessment of standards and certification. Toward sustainability: the roles and limitations of certification.* RESOLVE, Inc. Washington, DC.
- Cashore, B., Auld, G. and Newsom, D. 2004 Governing through markets: forest certification and the emergence of non-state authority. Yale University Press, New Haven, US.
- Cashore, B., Gale, F., Meidinger, E. and Newsom, D. 2006 Confronting sustainability: forest certification in developing and transitioning countries. Faculty of Forestry and Environmental Studies Publication Series, Yale University, New Haven, USA.
- Cattaneo, A. 2001 Deforestation in the Brazilian Amazon: comparing the impacts of macroeconomic shocks, land tenure and technological change. *Land Economics* 77: 291–240.
- Cerutti, P.O., Tacconi, L., Lescuyer, G. and Nasi, R. 2012 Cameroon's hidden harvest: commercial chainsaw logging, corruption and livelihoods. *Society and Natural Resources: An International Journal*, DOI:10.1080/08941920.2012.714846.
- Cerutti, P.O., Tacconi, L., Nasi, R. and Lescuyer, G. 2011 Legal vs. certified timber: preliminary impacts of forest certification in Cameroon. *Forest Policy and Economics* 13: 184–190.
- Chambers, R., 2009 So that the poor count more: using participatory methods for impact evaluation. *Journal of Development Effectiveness* 1: 243–246.
- Chen, H.T., Donaldson, S.I. and Mark, M.M. 2011a Validity frameworks for outcome evaluation. *New Directions for Evaluation* 130: 5–17.
- Chen, J., Tikina, A., Kozak, R., Innes, J., Duinker, P. and Larson, B. 2011b The efficacy of forest certification: perceptions of Canadian forest products retailers. *The Forestry Chronicle* 87: 636–643.
- Chen, J. M., Innes, J.L. and Tikina, A. 2010 Private cost-benefits of voluntary forest product certification. *International Forestry Review* 12: 1–12.
- Clark, M.R. and Kozar, J.S. 2011 Comparing sustainable forest management certifications standards: a meta-analysis. *Ecology and*

- Society. 16: 3 www.ecologyandsociety.org/vol16/iss1/art3/.
- Coleman, E.A. 2009 Institutional factors affecting ecological outcomes in forest management. *Journal of Policy Analysis and Management* 28: 122–146.
- Coleman, E.A., and Steed, B.C. 2009 Monitoring and sanctioning in the commons: an application to forestry. *Ecological Economics* 68: 2106–2113.
- Collier, D. 2011 Understanding process tracing. *Political Science and Politics* 44: 823–830.
- Cook, T.D., Scriven, M., Coryn, C.L.S. and Evergreen, S.D.H. 2010 Contemporary thinking about causation in evaluation: a dialogue with Tom Cook and Michael Scriven. *American Journal of Evaluation* 31: 105–117.
- Corso, E., DeSmet, S., Fernandez, I., Harrison, D., Portinga, A. and Woelders, L. 2008 Literature study and comparative analysis of the benefits of FSC certification in community forestry. Wageningen University and Research Centre, Wageningen, The Netherlands.
- Crosse, W., Newsom, D. and Kennedy, E.W. 2012 Recommendations for improving research on certification impacts (Appendix I). *In: Steering committee of the state-of-knowledge assessment of standards and certification. Toward sustainability: the roles and limitations of certification.* Washington, DC: RESOLVE, Inc.
- Crow, S. and Danks, C. 2010 Why certify? Motivations, outcomes and the importance of facilitating organizations in certification of community-based forestry initiatives. *Small Scale Forestry* 9: 195–211.
- Cubbage, F., Diaz, D. and Yapura, P. 2010 Impacts of forest management certification in Argentina and Chile. *Forest Policy and Economics* 12: 497–504.
- Cubbage, D.W., Richter, D., Thompson, M., Moore, S., Edeburn, J., Chesnutt, M., Cox, J. and Boyette, W. 2003 Forest certification costs and benefits: the experience of the Southern Center for Sustainable Forests. Washington, DC. Society of American Foresters. <http://scsf.env.duke.edu/files/CertNCsafprocs.pdf>.
- Damette, O. and Delacote, P. 2011 Unsustainable timber harvesting, deforestation and the role of certification. *Ecological Economics* 70: 1211–1219.
- Ebeling, J. and Yasué, M. 2009 The effectiveness of market-based conservation in the tropics: forest certification in Ecuador and Bolivia. *Journal of Environmental Management* 90: 1145–1153.
- Eden, C. 2009 The work of environmental governance networks: traceability, credibility and certification by the Forest Stewardship Council. *Geoforum* 40: 383–394.
- Elliott, C. 2000. Forest certification: a policy network perspective. CIFOR, Bogor, Indonesia.
- Ferraro, P.J., Hanauer, M.M. and Sims, K.R.E. 2011 Conditions associated with protected area success in conservation and poverty reduction. *Proceedings of the National Academy of Sciences* 108: 13913–13918.
- Ferraro, P. 2009 Counterfactual thinking and impact evaluation in environmental policy. *New Directions for Evaluation* 122: 75–84.
- Ferraro, P.J. and Pattanayak, S. 2006 Money for nothing? a call for empirical evaluation of biodiversity conservation investments. *PLOS Biology* 4: 482–488.
- Fischer, C., Aguilar, F., Jawahar, P. and Sedjo, R. 2005 Forest certification; towards common standards? Discussion paper 05-10. Resources for the Future, Washington, DC.
- Forest Stewardship Council (FSC) 2013 www.fsc.org/fileadmin/web-data/public/document_center/national_FSC_standards/FSC_Approved_FSS.pdf
- Frost, B., Mayers, J. and Roberts, S. 2003 Growing credibility: the impact of forest certification on forests and people in South Africa. Forestry and Land Use Program. International Institute for Environment and Development (IIED), London.
- Fujita, N. 2010 Using systems concepts in evaluation. *In: Fujita, N. (ed). Beyond logframe: using systems concepts in evaluation.* Foundation for Advanced Studies on International Development, Tokyo, Japan.
- Furman, R. 2005 Lessons from implementing a theory-of-change M&E system in the Livelihoods and Landscapes Strategy initiative (draft version). Forest Conservation Program. International Union for Conservation of Nature (IUCN). Gland, Switzerland.
- Gale, F. 2006 The political economy of sustainable development: the Forest Stewardship Council experience. Refereed paper presented to the

- Second Oceanic Conference on International Studies University of Melbourne. 5–7 July. Melbourne, Australia.
- Gan, J. 2005 Forest certification costs and global forest product markets and trade: a general equilibrium analysis. *Canadian Journal of Forest Research* 35: 1731–1743.
- Gaveau, D.L.A., Curran, L.M., Paoli, G.D., Carlson, K.M., Wells, P., Besse-Rimba, A., Ratnasari, D. and Leader-Williams, N. 2012 Examining protected area effectiveness in Sumatra: importance of regulations governing unprotected lands. *Conservation Letters* 5: 142–148.
- Garbarino, S. and Holland, J. 2009 Quantitative and qualitative methods in impact evaluation and measuring results. Governance and Social Development Resource Centre (GSDRC). Birmingham, UK.
- Gertler, P.J., Martinez, S., Premand, P., Rawlings, L.B. and Vermeersch, C.M.J. 2011 Impact evaluation in practice. World Bank, Washington, DC. http://siteresources.worldbank.org/EXT/HDOFFICE/Resources/5485726-1295455628620/Impact_Evaluation_in_Practice.pdf.
- Government Accountability Office (GAO) 2009 Program evaluation: a variety of rigorous methods can help identify effective interventions in United States. Government Accountability Office, Washington, DC.
- Greenstone, M. and Gayer, T. 2007 Quasi experimental and experimental approaches to environmental economics. *The Frontiers of Environmental Economics – Resources for the Future*. Discussion Paper 07-22. Washington, DC.
- Guijt, I. and Retolaza, I. 2012 Defining theory-of-change. Hivos, Discussion Paper <http://www.hivos.nl/content/download/76473/665916/version/1/file/E-Dialogue+1+What+is+ToC+Thinking.pdf>
- Guijt, I., Brouwers, J., Kuster, C., Prins, E. and Zeylanova, B. 2011 Evaluation revisited. Improving the quality of evaluative practice by embracing complexity. Conference Report. Center for Development Innovation, Wageningen University. Wageningen, The Netherlands.
- Hartsfield, A. and Ostermeier, D. 2003 Certification. The view from FSC-certified land managers. *Journal of Forestry* 101: 32–36.
- Holling, C. S. 1978 Adaptive environmental assessment and management. John Wiley and Sons, New York.
- Holvoet, B. and Muys, M. 2003 Sustainable forest management worldwide: a comparative assessment of standards. *International Forestry Review* 6: 99–122.
- Humphries, S.S. and Kainer, K.A. 2006 Local perception of forest certification for community-based enterprises. *Forest Ecology and Management* 235: 30–43.
- Hyde, W. 2012 The global economics of forestry. Resources for the Future. RFF Press, Washington, DC.
- International Tropical Timber Organization (ITTO) 2005 Revised ITTO criteria and indicators for the sustainable management of tropical forests including reporting format. ITTO Policy Development Series 15. Yokohama, Japan.
- Jagger, P., Sills, E.O., Lawlor, K. and Sunderlin, W.D. 2010 A guide to learning about livelihoods impacts of REDD+ projects. CIFOR Occasional Paper 56. CIFOR, Bogor, Indonesia.
- James, C. 2011 Theory-of-change review. A report commissioned by Comic Relief, Comic Relief: <http://mande.co.uk/2012/uncategorized/comic-relief-theory-of-change-review>
- Johansson, J. and Gun, L. 2011 Can voluntary standards regulate forestry? Assessing the environmental impacts of forest certification in Sweden. *Forest Policy and Economics* 13: 191–198.
- Karmann, M. and Smith, A. (eds). 2009 FSC reflected in scientific and professional literature: literature study on the outcomes and impacts of FSC Certification. FSC International Center. Bonn, Germany.
- Keskitalo, E.C.H., Sandström, C., Tysiachniouk, M. and Johansson, J. 2009 Local consequences of applying international norms: differences in the application of forest certification in northern Sweden, northern Finland, and northwest Russia. *Ecology and Society* 14 <http://www.ecologyandsociety.org/vol14/iss2/art1/>
- Kleemann, L. and Abdulai, A. 2012 Organic certification, agro-ecological practices and return on investment: farm level evidence from Ghana. Kiel Working Paper 1816. Kiel Institute for the World Economy. Kiel, Germany.

- Kollert, W. and Lagan, P. 2007 Do certified tropical logs fetch a market premium? A comparative price analysis from Sabah, Malaysia. *Forest Policy and Economics* 9: 862–868.
- Laufer, J., Michalski, F. and Peres, C.A. 2013 Assessing sampling biases in logging impact studies in tropical forests. *Tropical Conservation Science* 6: 16–34. www.tropicalconservationscience.org
- Lee, K. and Norris, R. 2012 Toward more sustainable production and consumption (Appendix B) *In: Steering committee of the state-of-knowledge assessment of standards and certification. Toward sustainability: The roles and limitations of certification.* RESOLVE, Inc., Washington, DC.
- Levin, S., Xepapadeas, T., Crepin, A.-S, Norberg, J., de Zeeuw, A., Folke, C., Hughes, T.P. Arrow, K.J., Barrett, S., Daily, G., Ehrlich, P., Kautsky, N., Mäler, K.G., Polasky, S., Troell, M., Vincent, J. and Walker, B. H. 2013 Social-ecological systems as complex adaptive systems: modeling and policy implications. *Environment and Development Economics* 18: 111132. doi:10.1017/S1355770X12000460.
- Marx, A. and Cuypers, D. 2010 Forest certification as a global environmental governance tool: what is the macro-effectiveness of the Forest Stewardship Council? *Regulation & Governance* 4, 408–434.
- McDermott, C.L. 2012 Trust, legitimacy and power in forest certification: a case study of the FSC in British Columbia. *Geoforum* 43: 634–644.
- McDermott, C.L. In press. Certification and equity: applying an 'equity framework' to compare certification schemes across product sectors and scales. *Environmental Science and Policy*.
- McDermott, C.L., Noah, E. and Cashore, B. 2008 Differences that 'matter'? A framework for comparing environmental certification standards and government policies. *Journal of Environmental Policy & Planning* 10: 47–70.
- McElwee, P.D. 2010 Resource use among rural agricultural households near protected areas in Vietnam: the social costs of conservation and implications for enforcement. *Environmental Management* 45: 113–131.
- Mickwitz, P. and Birnbaum, M. 2009 Key insights for the design of environmental evaluations. *New Directions for Evaluation* 122: 105–113.
- Mollinga, P.P. 2010 Boundary work and the complexity of natural resources management. *Crop Science* 50: S1–S9.
- Molnar, A. 2004 Forest certification and communities. *International Forestry Review* 6: 173–179.
- Nebel, G., Quevedo, L., Jacobsen, J.B. and Helles, F. 2005 Developing and economic significance of forest certification: the case of FSC in Bolivia. *Forest Policy and Economics* 7: 175–186.
- Nelson, A. and Chomitz, K.M. 2011 Effectiveness of strict vs. multiple use protected areas in reducing tropical forest fires: a global analysis using matching methods. *Plos One* 6: e22722.
- Nelson, V. and Martin, A. 2011 Impact evaluation of social and environmental voluntary standard systems (SEVSS): using theories of change. Working Paper 1, November. Natural Resources Institute, University of Greenwich, UK.
- Newsom, D., Bahm, V. and Cashore, B. 2006 Does forest certification matter? An analysis of operation level changes required during the SmartWood certification process in the United States. *Forest Policy and Economics* 9: 197–208.
- Newsom, D. and Hewitt, D. 2005 The global impact of SmartWood certification. Rainforest Alliance, New York
- Newsom, D., Kennedy, E., Miller, J., Bahn, V. and Adhikari, S. 2012 Testing a BMP-based approach for assessing gaps in certification impacts research (Appendix H). *In: Steering committee of the state-of-knowledge assessment of standards and certification. Toward sustainability: the roles and limitations of certification.* RESOLVE, Inc., Washington, DC.
- Nittler, J.B. and Nash, D.W. 1999 The certification model for forestry in Bolivia. *Journal of Forestry* 97: 32–36.
- Nussbaum, R. and Simula, M. 2004 Forest certification: a review of impacts and assessment frameworks. The Forest Dialog. Yale University, School of Forestry and Environmental Studies. New Haven, US.
- Nussbaum, R., Garforth, M., Scrase, H. and Wenban-Smith, M. 2001 An analysis of current FSC: accreditation, certification and standard-setting procedures identifying elements which create constraints for small forest owners. Proforest. Oxford, UK. www.proforest.org

- dfid.gov.uk/r4d/PDF/Outputs/Forestry/R7589_Output1.pdf
- Organisation for Economic Co-operation and Development (OECD)/Development Assistance Committee (DAC) 2002 Glossary of key terms in evaluation and results based management. The DAC Working Party on Aid Evaluation, OECD, Paris. www.oecd.org/dataoecd/29/21/2754804.pdf
- Oliver, R. 2004. The certification matrix: finding your way through forest certification schemes. Confederation of European Paper Industries (CEPI). Brussels, Belgium. www.forestrycertification.info. [accessed May 2011]
- Overdevest, C. 2010 Comparing forest certification schemes: the case of ratcheting standards in the forest sector. *Socio-Economic Review* 8: 47–76.
- Pacheco, P. 2004 What lies behind decentralization? Forest, powers and actors in lowlands Bolivia. *European Journal of Development Research* 16: 90–109.
- Pattanayak, S.K., Wunder, S. and Ferraro, P.J. 2010 Show me the money: Do payments supply environmental services in developing countries? *Review of Environmental Economics and Policy* 4: 254–274.
- Peña-Claros, M., Blommerde, S. and Bongers, F. 2009 Assessing the progress made: an evaluation of forest management certification in the tropics. *Tropical Resource Management Papers* 95, Wageningen University, The Netherlands.
- Pereira, D., Santos, D., Vedoveto, M., Guimaraes, J. and Verissimo, A. 2010 *Fatos Florestais da Amazonia 2010*. IMAZON. Belem, Brazil.
- Pintér, L. Hardib, P., Martinuzzic, A., and Halla, J. 2012 Bellagio STAMP: principles for sustainability assessment and measurement. *Ecological Indicators* 17: 20–28.
- Platt, J.R. 1964 Strong inference – certain systematic methods of scientific thinking may produce much more rapid progress than others. *Science* 146: 347–352.
- Pokorny, B. and Steinbrenner, M. 2005 Collaborative monitoring of production and costs of timber harvest operations in the Brazilian Amazon. *Ecology and Society* 10: 3. www.ecologyandsociety.org/vol10/iss1/art3
- Prabhu, R., Colfer, C.J.P., Venkateswarlu, P., Tan, L.C., Soekmadi, R. and Wollenberg, E. 1998 Testing criteria and indicators for the sustainable management of forests: Phase 1 Final Report. CIFOR Special Publication. CIFOR, Bogor, Indonesia.
- Prokopovych, B. 2010 Why do FMUs voluntarily comply with sustainability standards? The drivers and enabling conditions of voluntary certification (Appendix I). *In: Steering committee of the state-of-knowledge assessment of standards and certification. Toward sustainability: the roles and limitations of certification*. RESOLVE, Inc., Washington, DC.
- Pullin, A.S., Knight, T.M. and Watkinson, A.R. 2009 Linking reductionist science and holistic policy using systematic reviews: unpacking environmental policy questions to construct an evidence-based framework. *Journal of Applied Ecology* 46: 970–975.
- Putz, F.E. 2004 Multiple quests for the best criteria and indicators of sustainable tropical forest management: Worthwhile endeavors or a smokescreen? *Ecology and Society* www.ecologyandsociety.org/vol9/iss1/resp1
- Rametsteiner, E. and Simula, M. 2003 Forest certification – an instrument to promote sustainable forest management? *Journal of Environmental Management* 67: 87–98.
- Retolaza, I. 2011 Theory-of-change: a thinking and action approach to navigate in the complexity of social change processes. *Hivos/UNDP/Democratic Dialogue*. The Hague, The Netherlands.
- Richards, M. 2004 Certification in complex socio-political settings: looking forward to the next decade. *Forest Trends* www.forest-trends.org/documents/index.php?pubID=125
- Rivera, J. 2002 Assessing a voluntary environmental initiative in the developing world: the Costa Rican certification for sustainable tourism. *Policy Sciences* 35: 333–360.
- Rivera, J. and de Leon, P. 2004 Is greener whiter? Voluntary environmental performance of western ski areas. *The Policy Studies Journal* 32: 417–437.
- Roberts, T.M. 2012 Innovations in governance: a functional typology of private governance institutions (Appendix L). *In: Steering committee of the state-of-knowledge assessment of standards and certification. Toward sustainability: the roles and limitations of certification*. RESOLVE, Inc., Washington, DC.

- Rogers, P. J. 2008 Using programme theory to evaluate complicated and complex aspects of interventions. *Evaluation* 14: 29–48.
- Rogers, P.J. 2009 Matching impact evaluation design to the nature of the intervention and the purpose of the evaluation. *Journal of Development Effectiveness* 1: 217–226.
- Rogers, P. 2012 Introduction to impact evaluation. *Impact Evaluation Notes #1*. InterAction: A united voice for Global People and Rockefeller Foundation, New York.
- Romero, C. and Castrén, T. 2013 Approaches to measuring the conservation impact of forest management certification. Working Paper. Program on Forests (PROFOR), Washington, DC.
- Rosenbaum, P.R. and Rubin, D. 1983 The central role of the propensity score in observational studies for causal effects. *Biometrika* 70: 41–55.
- Rossi, P.H., Lipsey, M.W. and Freeman, H. 2004 *Evaluation: a systematic approach*. Sage Publications, Thousand Oaks, US.
- Rowe, A.E. 2012 Getting past the stale RCT debate: a new approach for impact evaluation. South African Monitoring and Evaluation Association Virtual Conference www.wpeg.wits.ac.za
- Ruben, R. and Zúñiga, G. 2010 How standards compete: comparative impact of coffee certification schemes in Northern Nicaragua. *Supply Chain Management: An International Journal* 16: 98–109.
- Rugh, J., Steinke, M., Cousins, B. and Bamberger, M. 2010 Summary of discussion of the alternative to the statistical counterfactual think tank. American Economic Association (AEA) 2010 Annual Conference, San Antonio, US.
- Ruiz Perez, M., Ezzine de Blas, D., Nasi, R., Sayer, J. A., Sassen, M., Angoue, C., Gami, N., Ndoye, O., Ngono, G., Nguingui, J-C., Nzala, D., Toirambe, B. and Yalibanda, Y. 2005 Logging in the Congo Basin: a multi-country characterization of timber companies. *Forest Ecology and Management* 214: 221–236.
- Ruslandi, R., Klassen, A., Romero, C. and Putz, F.E. (In press). Forest Stewardship Council certification of natural forest management in Indonesia: required improvements, costs, incentives and barriers. *Forests Under Pressure: Local Responses to Global Issues*. International Union of Forest Research Organizations (IUFRO). Vienna, Austria.
- Salazar, M. and Gretzinger, S. 2005 Costos y beneficios de la certificación forestal y mecanismos para la resolución de obstáculos comunes. WWF Centroamérica and PROARCA. San José, Costa Rica.
- Schreckenberg, K., Camargo, I., Withnall, K., Corrigan, C., Franks, P., Roe, D., Scherl, L.M. and Richardson, V. 2010 Social assessment of conservation initiatives: a review of rapid methodologies. *Natural Resource Issues* 22. International Institute for Environment and Development (IIED). London.
- Schulte-Herbruggen, B. and Davies, G. 2006 *Wildlife conservation and tropical timber certification*. Zoological Society of London, London.
- Steering committee of the state-of-knowledge assessment of standards and certification (SCR). 2012. *Toward sustainability: the roles and limitations of certification*. RESOLVE, Inc., Washington, DC.
- Shahwahid H.O. 2006 Timber certification in Malaysia. *In: Cashore et al.* (eds). *Confronting sustainability: forest certification in developing and transitioning countries*. Yale Faculty of Forestry and Environmental Studies Publication Series, New Haven, US.
- Sheil, D., Putz, F.E. and Zagt, R.J. (eds). 2010 *Biodiversity conservation in certified forests*. European Tropical Forest Research Network 51: 1–204.
- Sims, K.R.E. 2010 Conservation and development: evidence from Thai protected areas. *Journal of Environmental Economics and Management* 60: 94–114.
- Simula, M., Astana, S. and Ishmael, R. 2004 Report on financial cost-benefit analysis of forest certification and implementation of phased approaches. International Tropical Timber Council, Rep. ITTC (XXXVII)/13.
- Stein, D. and Valters, C. 2012 Understanding ‘theory-of-change’ in international development: a review of existing knowledge. Asia Foundation and the Justice and the Security Research Programme, London.
- Stern, E., Stame, N., Mayne, J., Forss, K., Davies, R. and Befani, B. 2012 Broadening the range of designs and methods for impact evaluation. Working Paper 38. Department for International Development (DFID), London.
- Sunderlin, W.D., Hatcher, J. and Liddle, M. 2008 From exclusion to ownership? Challenges and opportunities in advancing forest tenure reform. Rights and Resources Initiative, Washington, DC.

- Takahashi, T., van Kooten, G.C. and Vertinski, I. 2003 Why might forest companies certify? Results from a Canadian survey. *International Forestry Review* 5: 329–337.
- Tikina, A. and Innes, J.L. 2008 A framework for assessing the effectiveness of forest certification. *Canadian Journal of Forest Research* 38: 1357–1365.
- Tind, S. 2012 Main report – Assessment of certification and legality verification schemes for European Timber Trade Federation (ETTF), Almere, The Netherlands.
- Ton, G. 2012 The mixing of methods: a three-step process for improving rigour in impact evaluations. *Evaluation* 18: 5–25.
- Treves, A. and Jones, S. 2010 Strategic tradeoffs for wildlife-friendly eco-labels. *Frontiers in Ecology and the Environment* 8: 491–498.
- van Kooten, G.C., Nelson, H.W. and Vertinsky, I. 2005 Certification of sustainable forest management practices: a global perspective of why countries certify. *Forest Policy and Economics* 7: 857–867.
- van Kuijk, M., Putz, F.E. and Zagt, R. 2009 Effects of forest certification on biodiversity. Tropenbos International. Wageningen, The Netherlands.
- van Stolk, C., Ling, T. and Reding, A. 2011 Monitoring and evaluation in stabilisation interventions: reviewing the state of the art and suggesting ways forward. RAND Europe, prepared for Department for International Development (DFID) Stabilisation Unit, London.
- Viana, V., Ervin, J., Donovan, R.Z., Elliott, C. and Gholz, H (eds). 1996. Certification of forest products: issues and perspectives. Island Press, Washington, DC.
- Vidal, N. and Kozak, R. 2008 The recent evolution of corporate responsibility practices in the forestry sector. *International Forestry Review* 10: 1–13.
- Vogel, D. 2008. Private global business regulation. *Annual Review of Political Science* 11: 261–282.
- Vogel, I. 2012 Review of the use of ‘theory-of-change’ in international development. Review report. Department for International Development (DFID), London.
- Vogt, K., Larson, B., Gordon, J.C., Vogt, D.J. and Fanzeres, A. (eds.) 2000 Forest certification: roots, issue challenges and benefits. CRC Press, Boca Raton, US.
- Walrecht, A., Basso, K. and Hime, S. 2012 Certification and biodiversity: exploring improvements in the effectiveness of certification schemes on biodiversity. Climate Change and Sustainability Services. KPMG, Amsterdam, The Netherlands.
- Washburn, M.P. and Block, N. 2001 Comparing forest management certification systems and the Montreal Process criteria and indicators. Yale University School of Forestry and Environmental Studies and Pinchot Institute for Conservation. www.sustainableforests.net/docs/CI_Certification.pdf
- Weiss, C.H. 1997. Theory-based evaluation: past, present and future. *New Directions for Evaluation* 76: 41–55.
- West, S.G., Duan, N., Pequegnat, W., Gaist, P., Des Jarlais, D.C., Holtgrave, D., Szapocznik, J., Fishbein, M., Rapkin, B., Clatts, M., and Mullen, P.D. 2008. Alternatives to the randomized controlled trial. *American Journal of Public Health* 98: 1359–1366.
- White, H. 2009 Some reflections on current debates in impact evaluation. 3ie Working Paper 1. International Initiative for Impact Evaluation, London.
- White, H. 2010 A contribution to current debates in impact evaluation. *Evaluation* 16: 153–164.
- White, H. and Phillips, D. 2012 Addressing attribution of cause and effect in small impact evaluations: towards an integrated framework. 3ie Working Paper 15. International Initiative for Impact Evaluation, London.
- Wigboldus, S. and Brouwers, J. 2011 Rigid plan or vague vision: how precise does a theory-of-change need to be? Hivos E-dialogues. www.hivos.nl/eng/HivosKnowledgeProgramme/Themes/Theory-of-hange/Edialogues/E-dialogue-2
- WWF/World Bank Global Forest Alliance. Forest Certification Assessment Guide (FCAG) 2006 A framework for assessing credible forest certification systems/schemes. World Bank Global Forest Alliance, Washington, DC.
- Wunder, S. 2005 Macroeconomic change, competitiveness and timber production: a five country comparison. *World Development* 33: 65–86.
- Zagt, R., Sheil, D. and Putz, F.E. 2010 Biodiversity conservation in certified forests: an overview. *In*: Sheil, D., Putz, F.E. and Zagt, R. (eds). Biodiversity conservation in certified forests. ETRN and Tropenbos International. Wageningen, The Netherlands.

CIFOR Occasional Papers contain research results that are significant to tropical forest issues. This content has been peer reviewed internally and externally.

The often-claimed environmental and social benefits of forest certification remain to be empirically evaluated. Despite numerous publications on the impacts of tropical forest certification, virtually all are based on secondary sources of information and not on field-based measurements. This paper proposes an empirical research framework for a carefully designed field-based evaluation of the ecological, social, economic, and political impacts of tropical forest management certification taking into account location-specific contextual factors which shape certification outcomes. The paper also suggests that solid methodological quantitative and qualitative approaches be used to build proper counterfactuals on which to base the comparisons for inferring impacts, all informed by a thorough theory-of-change and through processes that bring stakeholders together. The proposed research framework represents a first step towards the design and future implementation of evaluation research in the context of tropical forest certification on a global basis. It is hoped the research framework proposed contributes to learning from past mistakes, building on lessons learned and enhancing decision-making towards the maintenance of forest values over the long term, and for the benefit of society as a whole.



RESEARCH
PROGRAM ON
Forests, Trees and
Agroforestry

This research was carried out by CIFOR as part of the CGIAR Research Program on Forests, Trees and Agroforestry (CRP-FTA). This collaborative program aims to enhance the management and use of forests, agroforestry and tree genetic resources across the landscape from forests to farms. CIFOR leads CRP-FTA in partnership with Bioversity International, CIRAD, the International Center for Tropical Agriculture and the World Agroforestry Centre.

cifor.org

blog.cifor.org



Center for International Forestry Research (CIFOR)

CIFOR advances human well-being, environmental conservation and equity by conducting research to help shape policies and practices that affect forests in developing countries. CIFOR is a member of the CGIAR Consortium. Our headquarters are in Bogor, Indonesia, with offices in Asia, Africa and South America.

